

Towards better biosecurity

Slowly but surely, a key advisory committee is helping the scientific community act more responsibly when conducting and publishing biological research that could carry security risks.

At an open meeting last week, the US National Science Advisory Board for Biosecurity (NSABB) demonstrated that it has made welcome progress in its daunting mission to sort out the responsibilities of stakeholders in relation to biological research that might put national and global security at risk. And although the board answers to the US government, its proceedings are relevant everywhere.

For example, the board's working group on communications has proposed a checklist to be used by anybody confronted with research results whose dissemination might assist terrorists or state-run bioweapons labs. The questions themselves are straightforward (for example: "Is novel scientific information provided that could be intentionally misused to threaten plant or animal health?"). They address both the potential risks and benefits of openness, and invite the user to weigh up one against the other. That is hardly profound or original: what's new is the idea that such a checklist should itself be widely disseminated to raise awareness and to help peer reviewers, university administrators, students, government officials and experienced investigators examine research critically.

Those used to such issues know that biosecurity risks are perceived by most people to outweigh the benefits of publication only where direct weaponization techniques are involved. But the draft checklist also raises questions about presentation: might the results as written cause public misunderstanding or encourage sensationalism? If the answer is yes, the working group lists a menu of options either for improved presentation or, if the results themselves are deemed too risky for open publication, to decide against publication altogether or to issue them on a 'need to know' basis.

Sensitive issue

This last option raises the spectre of a category of information defined as 'sensitive but unclassified', first promoted in the biological context by the Bush administration following the attacks of 11 September 2001. The NSABB working group, not empowered to decide on the merits of such a concept, had to consider it as a potential outcome. But neither the NSABB nor anyone else has come close to working out exactly how to implement such a category in practice.

The communications working group also espouses a set of overarching principles of responsible dissemination. These reiterate the fundamental benefits of scientific openness, stating that results should be communicated as fully as possible, and that communication should only rarely be restricted. This may sound like a truism to readers of *Nature*, but the working group includes individuals from security and military backgrounds. The endorsement of openness as a fundamental principle is significant, in effect placing the burden of proof of potential risk on those who might restrict communication.

The NSABB working group rightly places the onus on investigators and their institutions to assess their own work throughout the

research process, rather than on government or journals (although many publishers will also play their part). Researchers and their institutions know the work best, and keeping decisions in local hands is most likely to yield sensible results. It also minimizes the opportunities for overly burdened bureaucracies to opt for the least risky option, as they may see it, of restricting dissemination.

However helpful the draft guidelines are (they have yet to be formally adopted by the NSABB), they can only achieve so much. Ultimately, results that represent a biosecurity risk will get submitted, and scientists whose papers are rejected by one journal may submit them for publication elsewhere. And the guidelines don't specifically speak to the informal communication of 'dual-use' results in non-moderated forums, such as conferences or web postings.

Taking responsibility

The best the NSABB can do is plant the idea in scientists' minds that they should always think carefully about their work and its security implications — an idea that has not taken root in parts of the scientific community, which fears restrictions to its freedom. The NSABB is showing that the best way to keep the pursuit of knowledge open and free is for researchers to exercise a demonstrable sense of responsibility. At times, this may require scientists to ask and answer difficult questions. But if scientists don't learn to do this for themselves, the hard questions will come instead from outside parties, as happened recently with a modelling paper on botulinum toxin, whose publication was delayed following concern from the US government (see www.pnas.org/cgi/content/full/102/28/9737).

The NSABB has not yet addressed how to help scientists who wish to explore the potential risks of their results, and how to resolve disagreements with government agencies, for example, about how to communicate the results. This guidance is needed. For now, scientists and institutions engaged in research whose outcomes may carry security risks should take steps to incorporate such guidelines into their working practices, once the NSABB's final recommendations are established.

Previous attempts to address the problem of communicating dual-use research have arrived at broad principles that don't go far enough to help either scientists or administrators. By exploring specific questions and criteria that could be used to identify dual-use research, and by detailing how to determine whether the communication of its results should be handled with extra care, the NSABB is doing a service to the whole scientific community across the world. Its progress may seem to come at a snail's pace, but the board must be commended for taking a sensible approach, and for maintaining a consensus in the process. ■

"The best way to keep the pursuit of knowledge open and free is for researchers to exercise a demonstrable sense of responsibility."



Shooting the messenger

The abolition of a science advisory board to the US government sends the wrong signal.

The Secretary of Energy Advisory Board (SEAB) is not the most charismatic or influential body to offer advice to the US federal government, and few people will even notice it has gone. Nonetheless its abolition, without any satisfactory replacement (see page 725), once again raises the spectre of the Bush administration's loathing for anything that resembles objective outside advice.

Before dispensing with SEAB's services, energy secretary Samuel Bodman said he liked to operate with fewer advisers. This was a curious statement from the standard-bearer of an energy policy that has been dogged, since early in President Bush's first term, by allegations that it was fixed during closed-door meetings between Vice-President Dick Cheney and oil-industry lobbyists.

The Bush administration has made no secret of its contempt for time-honoured Washington government practice. It was during the presidency of Richard Nixon that Congress passed several laws intended to open up the workings of government to public scrutiny. One such measure was the 1972 Federal Advisory Committee Act, which established a framework for advisory panels in which government-appointed experts would meet to discuss the issues and advise the government, in full public view.

SEAB was such a panel, made up in part of scientists and engineers. Notwithstanding the fact that the energy secretary can

put who he wants on it, and the drawback that painful truths will often be kept quiet, it managed to do some good work. Last July, for example, it produced a stinging report on duplication in the energy department's nuclear-weapons laboratories. The labs' powerful supporters, led by Senator Pete Domenici (Republican, New Mexico), duly did their best to bury the report. Perhaps no one will ever know whether they also played a role in Bodman's decision to bury SEAB itself.

Over the years, SEAB has managed to point out issues of waste and overlap in government laboratories that no one — department officials, Congress, contractors or staff — particularly wanted to see aired.

It has periodically lent public backing to the energy department's vital but unheralded role in supporting basic research in the physical sciences. Manned as it was by such luminaries as Burton Richter, the Nobel laureate and former director of the Stanford Linear Accelerator Center, and the energy economist Daniel Yergin, its output could hardly be dismissed as the work of a few renegades.

SEAB cost less than half-a-million dollars a year to run, yet it provided oversight for a department with an annual budget of more than \$20 billion. If Bodman disliked the direction the panel was taking, he could easily have incorporated people whose views were more to his liking; at least their names would be public, for all to see. The decision to abolish it instead shows contempt for the principle of transparency in government. ■

"If Bodman disliked the direction the panel was taking, he could easily have incorporated people whose views were more to his liking."

Rightful owners

Research into anthropological artefacts must acknowledge claims of prior ownership.

Those seeking to study archaeological specimens increasingly find the objects of their interest tied up in complex arguments over who owns them. But if museums engage openly and honestly with nations that claim ownership, they can reach compromises that will balance the potentially conflicting desires for reliable stewardship, fairness and access for both the public and researchers.

The latest episode to bring these issues to the fore concerns the stewardship of a large collection of Oceanic artefacts held by the de Young Museum in San Francisco. The collection, which includes masks, skulls and exotic carvings, could provide anthropologists with new insights into primitive peoples.

As we report on page 722, at least nine of the 400 artefacts initially donated seem to be the national cultural property of Papua New Guinea. The items — and others among the 6,000 pieces lined up as future exhibits — came from the worldwide trade in archaeological specimens that are increasingly valued as works of art.

The de Young Museum — rebuilt at great expense after earthquake damage — started the Oceanic exhibition last October to herald its reopening. Unfortunately, given the history of dealing in such objects, no one at the museum seems to have asked questions

about the provenance of the specimens when they first entered its gorgeous, copper-sheathed walls.

Yet to the museum's credit, its officials are now pledging to address the issue. A Papua New Guinea museum official has, for example, been invited to San Francisco next week to examine the artefacts. The museum's desire to work with the government of the objects' country of origin contrasts with the more haughty, colonial attitude that has sometimes characterized curators' approaches to this issue in the recent past.

The world of archaeology has been transfixed in recent months, for example, by a criminal trial in Italy of dealers and museum officials over looted artefacts that were sold on to the Metropolitan Museum of Art in New York and the Getty Center in Los Angeles.

The artefacts now housed in the de Young Museum are of considerable scientific interest. They offer anthropologists a window to cultures of an earlier time — possibly even to the era of the miniature hominids who lived to the west on the island of Flores a few thousand years earlier.

But such research should proceed on the basis of a firm and fair understanding between the museum that has physical custody of the objects and officials from their country of origin. All parties should work together to ensure that such an understanding is reached. ■

"Research should proceed on the basis of a firm and fair understanding between the museum and officials from the country of origin."

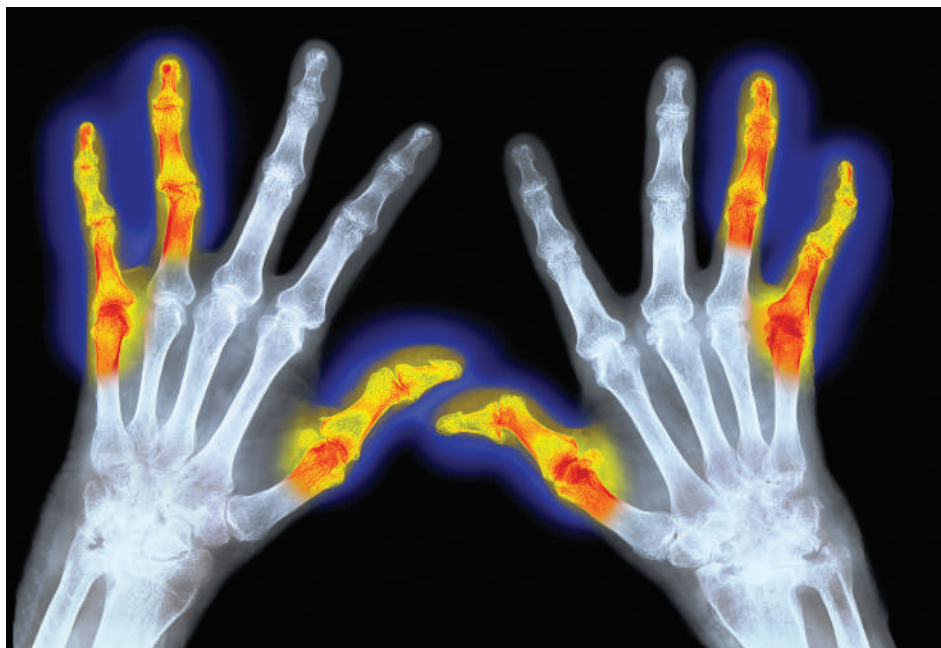
RESEARCH HIGHLIGHTS

Joint culprits

J. Clin. Invest. **116**, 961–973 (2006)

A newly discovered series of antibodies that attack the body play a central role in the autoimmune disorder rheumatoid arthritis, say researchers. The team suggests that future therapies that inhibit these antibodies may reduce or prevent the onset of the disease, which destroys joints (see X-ray, right).

Michael Holers of the University of Colorado Health Sciences Center, Denver, and his colleagues injected a cow-derived form of the protein collagen, known to induce arthritis, into mice. They found that antibodies targeting cyclic citrullinated peptides — proteins with modified arginine amino acids — were the first to appear before full onset of arthritis. Mice injected with the peptides and allowed to build a tolerance to them before the introduction of collagen contracted a less severe form of arthritis.



L. LEFKOWITZ/GETTY IMAGES

CANCER BIOLOGY

Tumour assassins

Science **311**, 1780–1784 (2006)

Although targeted therapies hold great potential for curing cancer, they often miss the mark because of inefficient delivery to the tumour. To improve their aim, Christopher Contag and his colleagues at Stanford University School of Medicine, California, combined two forms of these treatments.

The researchers started with a cytokine-induced killer T cell, a kind of immune cell that can home in on and kill cancer cells by detecting the abnormal proteins they make. The team boosted the T cell's killing power by inserting a tumour-killing pox virus into it. In tests on mice carrying human ovarian cancers or mouse breast cancers, whole-body imaging showed that the cells fully infiltrated the tumours and reduced them. Clinical trials utilizing this two-prong approach have been proposed.

ANIMAL BEHAVIOUR

Follow the herd

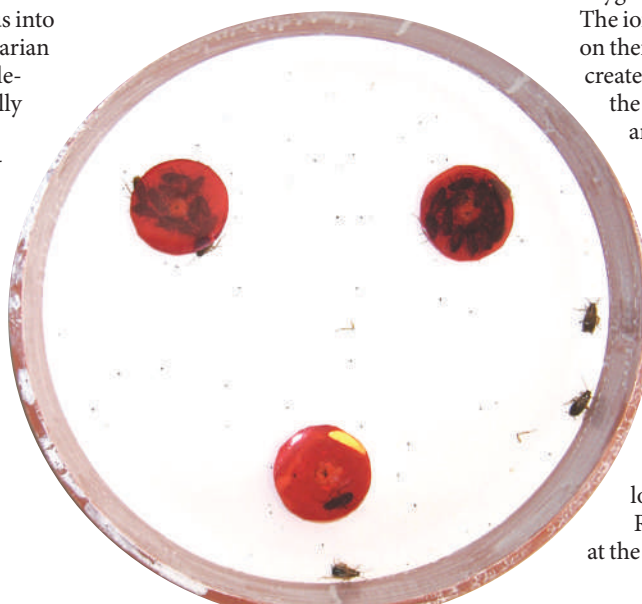
Proc. Natl Acad. Sci. USA

doi:10.1073/pnas.0507877103 (2006)

You don't need a big brain or a complex social structure to make smart group decisions, say biologists. A simple mathematical relationship between detecting a small amount of information and responding to it will do the job.

José Halloy at the Université Libre de Bruxelles, Belgium, and his colleagues tested cockroaches (*Blattella germanica*) to see how they choose to collect together in shelters. The decision to stay was influenced by the number of cockroaches already in the shelter, with the probability of a new arrival leaving for another shelter dropping sharply once a certain threshold of individuals had been reached, but only if there was sufficient room.

This resulted in optimally sized groups appearing in the shelters (as seen in the top two pictured below). The team suggests that similar rules could apply to other gregarious creatures such as shoals of fish and flocks of birds.



J. HALLOY

PHOTONICS

Silicon goes underground

Appl. Phys. Lett. **88**, 121108 (2006)

On silicon chips, space is at a premium. So Tejaswi Indukuri and his colleagues at the University of California, Los Angeles, have developed a method for stacking microstructures on top of one another, burying one below the surface of a silicon wafer while placing the other on top. In this way they have combined a photonic and an electronic device on the same area of a wafer.

They created a photonic device about 350 nanometres below the wafer surface by firing oxygen ions at it through a patterned mask. The ions penetrate to a depth that depends on their energy, and subsequent annealing creates a layer of silicon dioxide, isolating the silicon above from that below and thus defining a buried silicon microstructure. The second device can then be added on top.

CHEMICAL BIOLOGY

Think zinc

ACS Chem. Biol. **1**, 103–111 (2006)

Zinc is vital for many biological functions, but surprisingly little is known about its concentrations inside cells. A new method to quantify free zinc ions in cells looks set to change that.

Richard Thompson and his colleagues at the University of Maryland School of

IMAGESTATE/ALAMY

Medicine in Baltimore used a biosensor based on an enzyme called carbonic anhydrase. The biosensor fluoresces in the presence of zinc but does not react to other common metal ions such as calcium or magnesium. However, it was hard to get the biosensor into cells. So the team attached a protein fragment known to make cells take up proteins, causing sufficient quantities of the biosensor to enter cells. This allowed them to measure the picomolar concentrations of zinc typically found in ordinary resting cell culture lines for the first time.

ATMOSPHERIC SCIENCE

In-flight ozone info

J. Geophys. Res. **111**, D05305 (2006)
Atmospheric ozone levels in the tropics are increasing as a result of industrial emissions, according to the results of a survey that hitches a ride on cruising commercial aircraft.

Herman Smit of the Research Centre Jülich, Germany, and his colleagues on the MOZAIC (measurement of ozone and water vapour by Airbus in-service aircraft) programme have compiled details of the levels of these compounds between 1994 and 2003. They found that man-made emissions have boosted ozone levels in the troposphere — the layer of atmosphere nearest Earth's surface and below the one where the ozone layer resides — by around one part per billion per year. Levels of water vapour, which has similar greenhouse properties to ozone, varies by a factor of 2.5 during the year — far more than expected given the modest annual temperature variation that occurs at these latitudes.



FLUID DYNAMICS

Short spurt

Phys. Rev. Lett. **96**, 124501 (2006)

When water splashes on to a hydrophobic solid surface (pictured above), the result can be an extremely fine vertical jet travelling up to 40 times faster than the falling droplet. Denis Bartolo at the Ecole Nationale Supérieure in Paris, France, and his colleagues capture the process in moving pictures.

They find that the jets arise from the collapse, within a few millionths of a second, of an air bubble created in the middle of a droplet on impact. The fastest jets occur at two separate impact speeds and correspond to two fundamentally different collapse modes.

Surprisingly, at intermediate velocities, the bubble does not burst at all. Such insights could make a splash in areas such as inkjet printing in which the clean deposition of drops is essential, the authors suggest.

CHEMISTRY

Summing up

J. Am. Chem. Soc. doi:10.1021/ja058564w (2006)

Chemists in Israel have constructed a chemical calculator that can add or subtract three inputs made of single units of information, or bits.

Abraham Shanzer and his colleagues at The Weizmann Institute of Science in Rehovot made their device by modifying previous two-input addition systems based on the molecule fluorescein. By manipulating the amount of acid and base in the fluorescein solution, the researchers can control the

input bits and set the system to perform addition or subtraction. The answer is obtained by measuring the frequency of light emitted by the solution.

The authors say the system is the first to integrate three-input addition with subtraction, and is a significant step towards a molecular-scale calculator.

PROTEIN STRUCTURE

Proteins, pronto

J. Am. Chem. Soc. doi:10.1021/ja0580791 (2006)

Nuclear magnetic resonance (NMR) offers a fast way to determine a protein's structure, but samples must be free of the contaminants that are often introduced during the protein's production in a living cell. Now researchers have found a quick and clean method to go from a gene to a protein NMR spectrum in under five hours. This could allow highly efficient screening of proteins to make sure they have been expressed properly.

A. J. Shaka and his colleagues at the University of California, Irvine, and the Invitrogen Corporation in Carlsbad, used a cell-free protein production system called Expressway that is manufactured by the company. Purification yields a mixture clean enough for a two-dimensional NMR that shows the structural relationships of hydrogen and nitrogen atoms in the molecule.

JOURNAL CLUB

H. Charles J. Godfray
NERC Centre for Population
Biology, Imperial College
London, UK

An insect-population biologist is fascinated by the latest episode in a microbial detective story.

The now routine use of molecular techniques has revealed the existence of a previously unsuspected menagerie of bacteria living inside aphids and playing a key role in their biology.

Disentangling the effects of the bacteria on their hosts makes an intriguing story, although one that is a little unsettling for those of us who did traditional zoology courses that ignored microbes.

Like nearly all aphids, the pea aphid *Acyrtosiphon pisum* owes its life to a symbiotic bacterium, *Buchnera*, which provides essential amino acids that are absent from the aphid's diet. But aphids also have secondary symbionts that might be called 'lifestyle bacteria', which although not essential for life, can give the aphid an upper

hand in the struggle for existence.

At least five different types of bacteria have been associated with pea aphids, and a series of studies has shown that different secondary symbionts may help them to combat heat stress, exploit particular host plants and, perhaps most surprisingly, help them resist attack from parasitoid wasps and fungi.

But if secondary symbionts are so good for you, why do the majority of individuals carry just one and not the complete set? Researchers (K. M. Oliver *et al.*

Proc. R. Soc. Lond. B doi: 10.1098/rspb.2005.3436; 2006)
artificially created double infections using two secondary symbionts that they had previously shown increased resistance to parasitoids.

The doubly infected aphids had even higher resistance, but produced far fewer offspring, whether parasitized or not. Quite why double infections so harm their host is not clear, but the study provides an important insight into the miniature bacterial ecosystems at work inside aphids.

NEWS

Further accusations rock Japanese RNA laboratory

There is a “high possibility” that fraud was involved in a series of RNA studies from the laboratory of biochemist Kazunari Taira, a year-long investigation by the University of Tokyo has found. The papers in question “had no reproducibility and no credibility”, according to a statement dated 29 March.

The committee that carried out the inquiry also claims that a researcher on Taira's team may have faked data for the investigators, hoping to convince them that the results were reproducible. “Raw data submitted to the committee were fake,” says its chair, Yoichiro Matsumoto.

Co-author Hiroaki Kawasaki allegedly printed data for the committee obtained using software that was not available in 2003, when the experiment was conducted.

Kawasaki told *Nature* that he had possessed a demonstration copy of the software. The software maker, Applied Biosystems, told both the committee and *Nature* that no demonstration copy was available at the time.

Tipped off by a whistleblower in Taira's lab, the committee also tracked down a record that Kawasaki bought an enzyme he claimed to have constructed himself. Kawasaki says he bought it as a control when repeating his experiment. “I've never done anything wrong,” he says.

Taira has asked for retractions of four papers in question; he had already requested retraction of an earlier *Nature* paper¹. So far, Kawasaki

has not agreed to sign off on the retractions.

Taira's laboratory has developed novel RNA technologies to identify genes related to cancer and other diseases, preventing them from being expressed. In one paper², he and his colleagues reported that a human enzyme was expressed in the bacterium *Escherichia coli*, and this could be useful in making small fragments of RNA to suppress genes.

A second paper³ explained that introduction of small RNAs can induce DNA methylation — a process involved in some diseases such as cancer — in mammalian cells. The phenomenon had been reported in plant cells. The paper was later corrected, but not retracted.

On 26 March, Taira requested retraction of these two studies, along with two others^{4,5}. In his statement to the committee, Taira says he requested the retractions because of “a problem of research ethics”; and that Kawasaki had failed to provide lab notebooks to support the original results. *Nature* is deliberating on the retraction request, says editor-in-chief Philip Campbell.

When contacted by *Nature*, Taira said he stood by everything in his statement. He has in the past denied any wrongdoing.

Questions were first raised in April 2005 when the RNA Society of Japan, noting complaints from researchers, asked the University of Tokyo's School of Engineering to investigate whether 12 of Taira's papers were reproducible.



A University of Tokyo investigation has found the possibility of wrongdoing by some of its researchers.

The university selected a few of the papers and asked Taira to submit raw data by September. He failed to do so, blaming Kawasaki for allegedly not keeping good notebooks and storing back-up data on a computer that broke. The committee extended the deadline to the end of March, but again no evidence of reproducibility emerged.

A separate investigation was undertaken by the National Institute of Advanced Industrial Science and Technology (AIST), where Taira had headed a gene-discovery centre. Misconduct in nine of the twelve papers, it reported on 3 March, “could not be ruled out”. One other paper was in the clear, and two were not scrutinized as they did not involve AIST research. Last week, the institute announced that it

Doubts over evolution block funding by Canadian agency

A Canadian federal agency has denied funding to a science-education researcher partly because of its doubts about the theory of evolution.

Brian Alters, director of the Evolution Education Research Centre at McGill University in Montreal, had proposed a study of the effects of the popularization of intelligent design — the idea that an intelligent creator shaped life — on Canadian students, teachers, parents, administrators and policy-makers.

At a public lecture on 29 March,

Alters revealed excerpts from the rejection letter he received from the Social Sciences and Humanities Research Council (SSHRC). The letter stated that, among its reasons for rejection, the committee felt there was inadequate “justification for the assumption in the proposal that the theory of evolution, and not intelligent-design theory, was correct”.

“It illustrates how the misunderstanding of evolution and intelligent design can go to all levels of Canadian society,”

says Alters. David Green, director of McGill University's Redpath natural-history museum, adds: “I was quite surprised that such an opinion could be tendered by a high-powered granting agency.”

The SSHRC is the top social-sciences funding agency in Canada. Spokeswoman Eva Schacherl says its funding decisions are based on the comments submitted by a committee of peer reviewers, and that the council cannot comment on the sentence in question.

“We rely on the expertise of our committees to make recommendations,” she says.

Susan Bennett, an English professor at the University of Calgary in Alberta and chair of the SSHRC committee, could not be reached for comment by the time *Nature* went to press.

The maximum value of Alters's requested grant was Can\$40,000 (US\$34,000). He has received funding from the SSHRC before; for instance, his centre was awarded a Can\$175,000-grant to study the understanding of



JTB PHOTO/ALAMY

would not renew Taira's contract, which ended on 31 March.

A separate committee, set up earlier this year at the University of Tokyo, interviewed Taira, Kawasaki and other lab members about the lab's management and the possibility of misconduct. It plans to report to the university's president by 10 April, and to pass the issue to the university's disciplinary panel, which will decide possible penalties in the next few months.

Ichiko Fuyuno

Additional reporting by David Cyranoski

1. Kawasaki, H. & Taira, K. *Nature* **423**, 838–842 (2003).
2. Kawasaki, H., Suyama, E., Iyo, M. & Taira, K. *Nucleic Acids Res.* **31**, 981–987 (2003).
3. Kawasaki, H. & Taira, K. *Nature* **431**, 211–217 (2004).
4. Kawasaki, H. et al. *Nature Biotechnol.* **20**, 376–380 (2002).
5. Kawasaki, H. & Taira, K. *EMBO Rep.* **3**, 443–450 (2002).

biological evolution in Islamic societies.

Jennifer Robinson, associate vice-principal for communications at McGill, says the university will ask the council to review its decision. "In our view it is a factual error," she says. "The theory of evolution is a well-established science, and intelligent design is a religious belief."

Philip Sadler, a board member of the centre and director of science education at the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts, is more philosophical. "If he was trying to answer the question as to whether all this popularization had had an impact, he just saved the government \$40,000," says Sadler. "He found the evidence without doing the study."

Hannah Hoag

Korean science powerhouse sends Nobel laureate packing

Perhaps the relationship never had a chance. Robert Laughlin, who shared the 1998 Nobel Prize in Physics for his work on quantum fluids, was hired in July 2004 to reform the Korea Advanced Institute of Science and Technology (KAIST) in Daejeon. Last week, the institute's board of trustees decided not to renew his two-year contract.

Laughlin, formerly of Stanford University in California, was recruited by the office of South Korea's president, Moo Hyun Roh, to help shake up the country's university system. The reforms he proposed were fairly standard from a US perspective: he advocated competitive hiring of faculty at market rates; expanding the undergraduate student body; and running preparatory undergraduate degrees in

medicine, law and business.

Many of KAIST's 400 or so professors did not agree with Laughlin's approach. Of 278 that took part in a vote, 89% said they did not want him to continue. Choon Sup Yoon, a condensed-matter physicist who heads the faculty association, says that Laughlin's plans to expand the university's professional schools and humanities departments would have undermined KAIST's traditional status as a science and technology powerhouse. "His philosophy is totally different from what KAIST is about," says Yoon.

Another KAIST researcher, who did not want to be named, says the problem was more to do with Laughlin's attitude. "The content was not that bad. But he made enemies, probably because of a lack of experience in running

an organization," he says.

Laughlin admits that he lacked experience, but says that was not the real problem. The point of his dismissal, he says, "was to discredit the policy of hiring outsiders".

It was a stormy end to a bitter affair in which Laughlin was even at times accused of abusing his vacation privileges. Some say it was inevitable. Chan-Mo Park, head of Pohang University of Science and Technology where Laughlin still holds a post as distinguished professor, says that philosophies simply clashed. "Everything resulted from a cultural difference between the two countries," says Park. "It's a very sad story."

Laughlin finishes at KAIST in mid-July and will return to Stanford to resume teaching in the autumn.

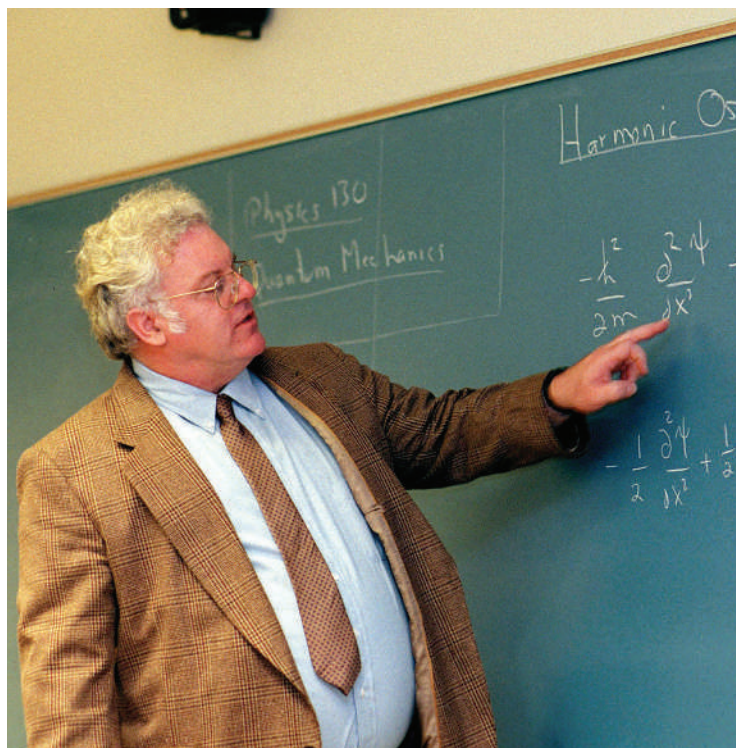
David Cyranoski



SOUTHERN INDIA SEES DROP IN HIV

Good news is tempered by continued global rise.
www.nature.com/news

N. FREDERIC/CORBIS SYGMA



By the board: Robert Laughlin's contract with a prestigious South Korean institute has not been renewed.

SPECIAL REPORT

Guinea experts cry foul on tribal exhibits

A new exhibition of Melanesian artefacts raises questions about how the pieces ended up in a Californian art museum. **Rex Dalton** investigates.

At San Francisco's new fine-arts museum, a stunning exhibition of South Pacific artefacts sheds light on the world of primitive peoples. The display includes about 400 specimens from Papua New Guinea — with adorned human skulls from the days of head-hunting, masks and shields portraying tribal spirits, and life-size carvings steeped in fertility rites.

The de Young Museum offers scientists the first opportunity to study a rare collection of some 6,000 Melanesian artefacts. But the exhibition also engenders difficult questions — because at least nine of the artefacts seem to be the national cultural property of Papua New Guinea. Records show that these have been earmarked by the country's government as objects of cultural significance to the nation.

Several anthropologists, many of whom have worked in Papua New Guinea, claim that the specimens were illegally exported. And other artefacts in the exhibition can allegedly be traced back to worldwide trade networks, in which misappropriated specimens are sold to antiquities dealers.

In recent months, accusations and criminal charges have arisen against high-profile art dealers and museum leaders. An ongoing criminal trial in Italy has revealed that looted artefacts from ruins have been sold to fine-arts museums such as the Metropolitan Museum of Art in New York and the Getty Center in Los Angeles. And some scholars at New York University are protesting over a \$200-million donation to the university, to fund a centre for the study of the ancient world; they say the funds may be tainted by antiquity dealing.

Concerns over the source of the de Young artefacts raise complex issues for the museum, recently rebuilt after it was damaged by an earthquake. Only 30 years independent, Papua New Guinea remains a poor nation, with little international clout or economic means to fight for its lost patrimony. Some question whether

the country's Public Museum and Art Gallery, in Port Moresby, would have enough money to properly secure and care for the artefacts if they were returned.

But others are adamant that the issue must be addressed. A 1965 law bans exports of national cultural property, says Barry Craig, curator of foreign ethnology at the Museum of South Australia in Adelaide. "Material removed illegally should be returned. But it is up to the Papua New Guinea government," argues Craig, who served as a curator at Papua New Guinea's national museum from 1980 to 1983.

The Papua New Guinea government has tried to crack down on illegal exports in the past. In 1972, working with the government, Brian Egloff, a University of Canberra ethnoarchaeologist, and Dirk Smidt, a Dutch ethnographer, shut down exports one weekend and discovered crates of prohibited artefacts ready for shipping abroad. Some were to go to Wayne Heathcote, an Australian collector, who has clashed with Papua New Guinea authorities over artefacts; Heathcote declined interview requests.

"All those artefacts that were removed illegally should go back."



But other artefacts have slipped out to auction or parlour deals, which can be lucrative; the donor of the de Young collection values his gift at \$100 million.

Finding the source

Smidt, who has long championed Papua New Guinea patrimony, penned an article in the de Young catalogue on the current exhibition. Records show that he had previously entered several of the nine artefacts into Papua New Guinea's national-property catalogue. Smidt



The de Young Museum offers scientists a rare chance to study a large collection of Melanesian artefacts.



Whose heritage? The carvings shown here have been classified as property of Papua New Guinea.

acknowledged that national property was in the collection, but declined further comment.

Egloff, who has investigated questionable artefact purchases in Australia, says an independent audit should be conducted on the de Young specimens. "There are two issues," says Egloff. "First, to understand the objects' place in society and how they came to the collection. And second, what to do with them."

Sebastine Haraha, the enforcement officer for the Port Moresby museum, says he will inspect the artefacts in San Francisco next week, in order to confirm that they are national property. Museum files contain notes claiming Heathcote illegally purchased two of the artefacts in the 1970s, Haraha says. "I expect all those that were removed illegally should go back."

Officials at the de Young say they are taking the issue seriously. The museum's director, art historian John Buchanan, says he has begun an inquiry into the origin of the questioned artefacts. The de Young museum has a history of repatriating objects when appropriate, Buchanan says, and could do it again.

"We want to be a positive player to resolve the puzzle," says Buchanan. "The way to do that is to be as transparent as we can."

The nine objects in question include a

Karawari mask from the Middle Sepik River (pictured, near left); a wooden Garamut slit gong, about three metres long; and a metre-tall carving showing a Kaningara woman copulating, preying-mantis style (pictured, far left).

Such unique artefacts cry out for study, but not as pieces of art, says anthropologist Adrienne Kaeppler, Oceania curator at the Smithsonian Institution in Washington DC. "These pieces need to be studied in conjunction with people knowledgeable about the culture from which they were taken."

Homeward bound?

Already, researchers at the University of Arizona in Tucson have conducted radiocarbon analysis on 150 specimens. More than 30 are at least 300 years old — with one mask dating to 1,200 years ago.

But the questioned provenance of some specimens may hamper research on the collection. One Smithsonian scientist says privately, "We wouldn't go near it now."

The Melanesian collection was donated by John and Marcia Friede. John Friede, who began collecting artefacts about 40 years ago, has purchased former European museum collections, expropriated in the late 1800s by missionaries. And for decades, he bought specimens from international dealers, such as Heathcote.

Friede denies knowing any of the Melanesian artefacts were national cultural property, insisting that he did not engage in any illegal exports. "Dealers don't tell you where they get them," he says, adding: "It really doesn't matter a great deal." Repatriating the artefacts is "crazy", he says. "They are a lot better off in San Francisco."

If international law-enforcement authorities confirmed that certain pieces were stolen, Friede says, he could see them being returned. "If a few specimens cause the collection to be thought of as nasty, illegal stuff, it would be a shame. My desire is to create a collection to define the greatness of the art."

Papua New Guinea's ambassador to the United States, Evan Paki, expressed surprise when *Nature* pointed out that the de Young collection probably contained national property, saying: "I'm not aware of that. I don't want to set something in motion. But if something is clearly protected, we would want it returned."

And in Papua New Guinea, the long wait for the repatriation of national property continues. As Prime Minister Michael Somare wrote in 1974: "We view our masks and art as living spirits with fixed abodes. It is not right they should be stored in New York, Paris, Bonn or elsewhere."



BETTER THAN X-RAYS
Shoebox-sized scanner is ideally suited to spotting drug caches
www.nature.com/news

Energy secretary ditches science advisers

WASHINGTON DC

The US Secretary of Energy, Samuel Bodman, has decided to disband the Department of Energy's principal independent advisory board on scientific and technical matters.

The Secretary of Energy Advisory Board (SEAB), which has existed in one form or another since 1978, will end after a final meeting this spring. Bodman made the decision about a month ago in a closed-door meeting with senior staff, says department spokesman Craig Stevens.

Some observers say that the decision is another example of the Bush administration's reluctance to accept outside scientific advice (see page 716). "This decision is part of a bigger picture," says Stephen Dean, president of Fusion Power Associates, a non-profit organization based in Gaithersburg, Maryland, that promotes fusion research. "This administration doesn't want outside advisory groups to tell it what to do."

But Stevens says that the choice has more to do with the scientific background of Bodman, who is a chemical engineer by training, than a desire to limit scientific advice. "The secretary has an understanding of science and scientific processes," Stevens says. "He would rather hear firsthand from the folks in the department." Stevens adds that plenty of direction is provided by the Advanced Energy and American Competitiveness Initiatives, which are intended to boost work on energy and basic research.

The 28-member, politically appointed SEAB



Stand alone: President Bush's energy department is headed by chemical engineer Samuel Bodman (left).

"The secretary would rather hear firsthand from the folks in the department."

is a mix of distinguished scientists, such as Nobel laureate Burton Richter, and business executives, such as former ExxonMobil chairman Lee Raymond. But not all of its reports have been influential, says Mark Marin, a lobbyist with Lewis-Burke Associates in Washington DC. One recent report, on research funding in the department, was incorporated into the president's competitiveness initiative, but a study on laboratory contracting practices was largely ignored.

The panel's most recent report, in July 2005,

recommended drastic restructuring of the nation's nuclear-weapons labs (see *Nature* 436, 316; 2005). The study riled some in Congress, but Stevens denies that this had any influence on the decision to dissolve the board.

The SEAB was not the only group to give scientific advice to the secretary; several boards regularly report on issues such as nuclear physics and energy research. Richter, of the Stanford Linear Accelerator Center in Menlo Park, California, says that Bodman can consult whomever he wishes: "The secretary will get his advice where he wants to get his advice." ■

Geoff Brumfiel

Wellcome Trust fuelled bid to save British science treasure

The Wellcome Trust put its financial weight behind Britain's Royal Society last week in order to rescue a prized seventeenth-century manuscript from bidders at Bonhams auction house in London.

The society claims that the 520-page treasure, handwritten by the physicist Robert Hooke, was taken from its archives some 300 years ago, and put out a public plea to recover the papers (see *Nature* 439, 638-639; 2006). The privately funded Wellcome Trust pledged half of the £1 million (US\$1.7 million)

that the society used to negotiate an eleven-hour deal with the manuscript's vendors.

The papers document the society's meetings from 1661 to 1682, and feature many personal comments from Hooke, former secretary of the society and one of its most influential scientists. The manuscript was found in a private house in Hampshire last year. When it became apparent in February that the papers were to be auctioned off, the society appealed for benefactors.

The resulting £1 million came

from 146 separate pledges, 122 of which were from Royal Society fellows. Donations ranged from £5 to £500,000, although a society spokesman told *Nature* that the Wellcome Trust was not the only substantial donor.

"We are very grateful," says the Royal Society's president, Martin Rees. "We're delighted because the manuscript has gone to its natural place." The society hopes to take delivery of the document this month, and is planning to display it as part of a summer exhibition at

the society's London headquarters.

Historians will scrutinize the handwritten papers for discrepancies with the formal minutes of the meetings, which the society already has. The results may resolve several disputes, such as Hooke's claim that he was the original inventor of the spring-balance watch.

Clare Matterson, director of medicine, society and history at Wellcome, says the manuscript is "equivalent to a Shakespeare folio". ■ Michael Hopkin

From the front lines

As the H5N1 flu virus continues to sweep across the globe, researchers in some of the countries affected describe in their own words the political and scientific challenges that they face.

INDONESIA

Andrew Jeremijenko

Medical epidemiologist and former head of influenza at the US Naval Medical Research Unit No. 2 in Jakarta

With 30 human cases of avian flu since last July — 23 of whom have died — Indonesia is currently the world's avian-flu hotspot. Vietnam and Thailand have stemmed the flow of cases by controlling the disease in animals. But in Indonesia, the cases keep coming. Every human case is a potential stepping stone to a pandemic virus, and each case in Indonesia should therefore be considered both a national and an international emergency.

Controlling the disease in Indonesia is inherently difficult. The country has 240 million people — and 1.3 billion chickens — spread across some 6,000 inhabited islands. Political power in this fledgling democracy is also highly decentralized across the archipelago.

The virus is now endemic in many of the country's millions of backyard farms. One reason is the initial delays in taking action, with Indonesia only belatedly admitting to the presence of H5N1 in poultry in January 2004. It has probably been present since August 2003.

Surveillance and identification of H5N1 in animals and birds need to be substantially improved, as does clinical and epidemiological follow-up, if we are to get a handle on the science of what is happening in Indonesia. There is a need for strong coordination between those involved in public health and in animal health.

Indonesia is now promptly reporting suspected and confirmed human cases to the World Health Organization (WHO); however, it is probably not detecting all cases. The mortality rate is also extremely high, with 12 of the 13 cases this year dying. That gives a 92% mortality rate, far higher than the global 56% rate. We need to identify and address the causes, such as delays in detecting cases.

In particular, we need to compare the sequences of the virus in

human cases and in surrounding poultry, to spot any changes in the virus that may be occurring in humans. Indonesia has generally been good at sharing human samples with the international community, but urgent requests for data on animal viruses found in the vicinity of human cases have been ignored.

On the bright side, the country has developed a national control and preparedness strategy that was signed by President Susilo Bambang Yudhoyono on 13 March. National and international pressure must be maintained to ensure that these improvements are sustainable.

THAILAND

Les Sims

Consultant for the Food and Agriculture Organization (FAO)

The outlook for avian flu in Thailand is most encouraging. The country has experienced several waves of animal infection, but since a peak in summer 2004, the number of outbreaks has progressively decreased, and no new outbreaks have occurred in the past three months.

One must be prudent, however, about the extent to which the strategies behind this success can be exported to other countries with structurally different poultry industries. Most of Thailand's poultry is in commercial flocks,

in sharp contrast to Vietnam and China, which have many backyard farms.

That said, some of the notable innovations in the Thai programme should be considered elsewhere. In particular, it has used village-level disease-spotters as surveillance sentinels. The procedure began in September 2004, when the government launched a nationwide survey. This involved hundreds of thousands of inspectors carrying out house-to-house searches in an attempt to obtain reliable data on incidence and so target control measures more effectively.

INDIA

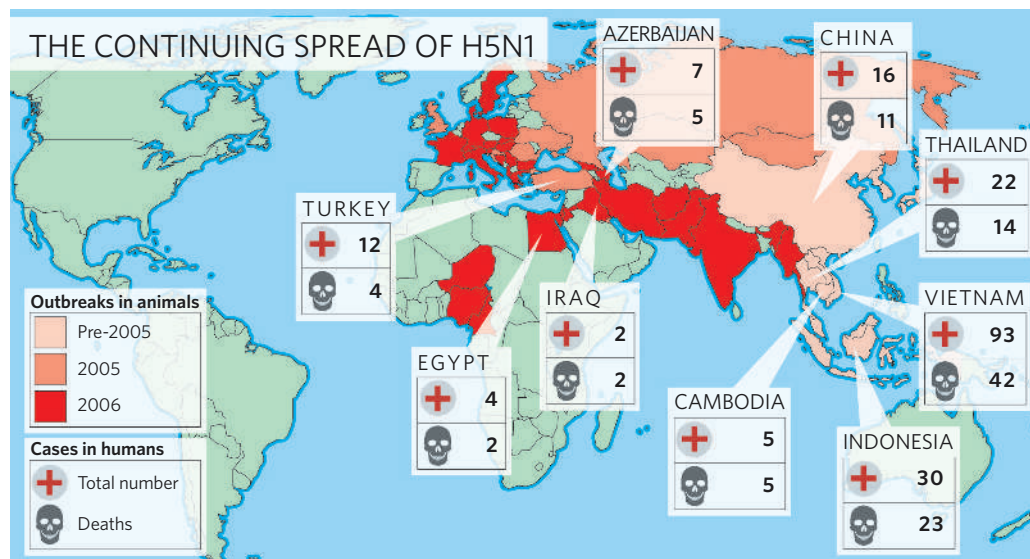
Shahid Jameel

Head of virology at the International Centre for Genetic Engineering and Biotechnology, New Delhi

Avian flu was first reported in India in February in the Navapur district of the western state of Maharashtra. The government response has been fair to good, with a low-key, yet aggressive strategy in which enough manpower, containment gear and medicines were deployed in affected areas.

Control is mainly by surveillance, and by culling birds in a 3-kilometre radius around outbreaks. One big weakness is the lack of laboratories that can carry out rapid surveillance. The entire country has only one facility with the containment security to handle bird H5N1 virus samples and only two that can process human samples.

Communication with the public has been a problem. At the same time that several government health officials ate chicken on live television to show it was safe, chicken was banned from the canteen in the Indian parliament and removed from the menus on state-run Indian Airlines and Indian Railways.



The High Security Animal Disease Laboratory in Bhopal is stretched. It is handling 4,000 samples a week and doesn't have the staff or capacity for this, which is leading to delays. Precious time is being lost in making decisions on where to cull, giving the virus the opportunity to spread. For example, positive results relating to an outbreak in Gujarat that began on 27 January were announced only on 16 March.

The major challenge is to keep the virus out of backyard flocks. And we need to think about vaccination and other strategies besides mass culling, as chicken and eggs are a cheap source of protein for much of the population.

NIGERIA

Claude P. Muller

Director of the Institute of Immunology, WHO Collaborative Center for Measles, Luxembourg

Avian flu broke out in commercial poultry farms in Kano and Kaduna states in northern Nigeria in January, but was misdiagnosed as Newcastle disease. H5N1 is at risk of becoming endemic in large parts of Africa, hitting already fragile food supplies and creating new opportunities for spread to humans. Despite the weakness of surveillance for avian flu in Africa, the international community has been very slow in reacting to its appearance on the continent.

We have helped set up the only laboratory in Nigeria that can diagnose H5N1, with Ademola Owoade, a poultry veterinary surgeon at the University of Ibadan. One lab is not enough for a country of 129 million people. The state organizes the collection of samples, and we supply reagents and staff.

Within a few days of setting the lab up, we found that the virus had already spread to the southwest, the hub of the country's poultry industry and far beyond the cordon sanitaire that the federal government had tried to set up in the north. It is probably safe to say that the virus is now endemic in the region.

The requirement of the World Organisation for Animal Health and the FAO — that samples be confirmed at reference laboratories before being officially acknowledged — is also allowing the virus time to spread. Although the economic impact of false positives may be great, false positives are less important than false negatives in terms of catching the virus before it spreads.

Poultry is the second-largest industry here after oil. Some farms have been very professional in taking comprehensive precautions, but many large farms have not (see www.aphis.usda.gov/birdbiosecurity/hpai.html). I believe that the virus cannot be contained in Africa without vaccination, as most farmers cannot afford to restock after culling their flocks.

AZERBAIJAN

Guenael Rodier

WHO special adviser who coordinated the response in Azerbaijan

Azerbaijan last month reported seven cases in humans, of whom five have died — its first cases of the disease. Worryingly, six of the cases were a cluster, occurring in the small Daikand settlement in the southeast. As there had been no obvious contact with poultry, this initially raised fears of possible human-to-human spread.

Now, however, we believe we have pinned down the cause: direct contamination from wild birds. It is not proven, but all six were involved in defeathering wild birds. Educating the public about high-risk practices, such as defeathering and slaughtering, will be our key task here.

CHINA

Guan Yi

Virologist at Hong Kong University

China has reported 16 human cases of H5N1 infection, 11 of them fatal. The latest case, reported in Shanghai on 24 March, is puzzling — like another recent case in Guangzhou — in that there were no reported poultry outbreaks nearby, and the victims had no close contact with poultry. We believe we may be looking at a route of contamination that has not been considered previously: direct and indirect contact with birds in live poultry markets.

Work by my group supports this hypothesis. Birds tested in markets in southern China over the past six months show that, despite extensive vaccination, only 20–50% of birds in the markets carry antibodies to H5N1. And H5N1 virus could be isolated from around 1% of the birds in the market, even though they appear healthy. They seem to have some partial protection, but are still shedding virus, as the rest of the poultry population might be doing. And when we test viruses isolated from market poultry in the lab, they are still highly pathogenic in experimental chickens. This means that either these birds have developed some natural immunity or there is a problem with the quality of the vaccines used. Current vaccination programmes have not prevented the virus from becoming prevalent in poultry populations in this region.

Looking back on five years of surveillance, it is also now clear that the incidence of H5N1 in domestic ducks and geese is much higher than in chickens, so these are the major vectors for H5N1. If we want to control H5N1 in China and southeast Asia, we need to control its spread in aquatic poultry.

Interviews by Declan Butler

ON THE RECORD

“Luckily criminals wear trainers. If they all wore Oxford brogues we would be in a very difficult position.”

Nigel Allison of the University of Sheffield describes a computerized system being developed in Britain to identify shoeprints at crime scenes.

“This Pope has remained a scientist at heart.”

Ernst-Ludwig Winnacker, president of Germany's main research funding agency, is impressed after meeting Pope Benedict XVI.

Sources: BBC, DFG

SCORECARD



Chinese health

Chinese authorities announce that they have wiped out lymphatic filariasis, a disease that can lead to grossly enlarged limbs. The World Health Organization hopes to eradicate the disease by 2020.



Ancient coiffure

Cleopatra employed up to three different hairstyles to present the best image to the locals wherever she travelled, says a recent archaeology book.



Climate change

The UK government admits that it will fail to meet its self-imposed goal of cutting carbon dioxide levels by 20% of 1990 levels by 2010.

NUMBER CRUNCH

ViaGen, a company in Austin, Texas, is cloning champion horses. The animals were all big in cutting, a sport that grew out of cattle roundups and involves singling out a cow from a herd. But what's at stake?

\$150,000 is how much ViaGen charges to clone a horse.

\$90,000 is how much the company charges for a second copy of the same horse.

\$380,000 is how much Royal Blue Boon, the first champion horse cloned, won in its sporting lifetime.

Sources: Washington Post, News8Austin

Fantasy reference list leads to the sack

What's in a name? Hui Liu, assistant dean of the medical school at Tsinghua University in Beijing, got his job after submitting a résumé that cited, among other things, a paper by "H. Liu" in the *Journal of General Virology*. He lost his job last month when the university learned that the paper was written by Hong Liu, a researcher at Mount Sinai School of Medicine in New York.

The school was told of the problem last November by Shi-min Fang, a San Diego-based biochemist whose home page is a popular place to post rumours of scientific fraud. Following up an anonymous tip-off, Fang says he investigated Liu's résumé, which had been posted on Tsinghua's website. "I found that several papers were either non-existent or belonged to somebody else," Fang says.

The university's investigation led to Hui's dismissal.

Japan tempts young lovers with science magazine

Japan's science ministry has unveiled its latest strategy to combat a declining interest in science among young people: a free magazine aimed at couples on hot dates.

Last week, the ministry issued 1.1 million copies of a 16-page pamphlet called *Science Walker* (pictured), which will be tucked into *Tokyo Walker* and distributed free at train stations and convenience stores.

The cover features a young couple, along with catchphrases such as "Science that couples can enjoy together". Topics for couples to discuss include the technology of a German football stadium, sports shoes with built-in sensors to check walking speed



Ancient optical illusion is easy on the eye

D. SILVERMAN/GETTY

This Roman mosaic is proving to be a hit with neuroscientists as well as archaeologists. Unearthed last November at Megiddo prison in Israel, the mosaic is on the floor of a Christian church dating back to the third or fourth century AD.

The diamonds in the corners of the image can be seen as three-dimensional cubes that either cut in to, or stand out of, the framed border, an effect known as ambiguous or bistable perception.

Similar optical illusions can be found in mosaics from various regions of the Roman Empire, but the Israeli find is a particularly fine example of the phenomenon, say cognitive



neuroscientists and art history buffs Eric Altschuler of the University of Medicine and Dentistry of New Jersey, Newark, and Alex Holcombe of Cardiff University, UK.

and ground condition, and science cafes where people can chat with leading scientists. Couples going out to dinner might enjoy stories about the high-tech freezing methods used to ensure that tuna at sushi restaurants is kept fresh and palatable.

The ministry is waiting for public feedback before deciding whether to continue with the pamphlets.

Researchers put a price on the value of insects

What are insects worth? How about \$57 billion a year in the United States alone?

Two conservation researchers compiled this estimate to stress the value of our six-legged friends. Although putting a cash value on the environment is difficult and controversial, ecologists do it to explain the worth of biodiversity to officials more used to dealing in dollars and cents.

The researchers totted up the economic transactions that would have been impossible without insects (J. E. Losey & M. Vaughan *BioScience* 56, 311–323; 2006). And \$57 billion is a conservative estimate, says Mace Vaughan of the Xerces Society for Invertebrate Conservation in Portland, Oregon, one of the paper's authors.

Some wild insects do cause problems, such as disease or crop damage, but overall their value is immeasurable, the authors say. And some of the costs, such as from crop pests, are mitigated by other insects.

Africa's crop crisis has roots in low-fertility soil

Three-quarters of the farmland in sub-Saharan Africa is severely degraded of nutrients, according to a report released

last week. The study, by Julio Henao and Carlos Baanante of the Alabama-based International Center for Soil Fertility and Agricultural Development, shows that some regions lose up to 60 kilograms of nutrients per hectare each year. As a result, cereal production in the region has stagnated at one tonne per hectare, compared with about three tonnes in the rest of the world.

The nutrient-starved soils are one of the major factors preventing the 200 million malnourished people in the continent from growing enough food to eat and sell, development experts said in New York last week. They will discuss possible solutions at the Africa Fertilizer Summit, to be held in Abuja, Nigeria, from 9–13 June.

Merger prompts change in oversight for Bell Labs

Bell Labs, the home of Nobel-winning physics research, may soon be under new ownership.

On 2 April, French telecommunications company Alcatel announced plans to purchase Bell Labs' parent company, Lucent Technologies, for US\$13.5 billion. The merger was approved unanimously by the boards of both companies.

Under the agreement, Bell Labs will continue to be based in Murray Hill, New Jersey (see *Nature* 440, 146–147; 2006). Sensitive research carried out for the government, in fields such as communication and cryptography, would come under the oversight of an independent board of US citizens, says Pat Russo, Lucent's chairwoman. The arrangement is meant to alleviate national-security concerns.

The merger could take 6–12 months to complete and must be approved by both the European Commission and US authorities.

The story of *i*

Multicellular creatures can be battlegrounds for competing populations of cells. **Claire Ainsworth** learns how this way of looking at an individual is feeding into immunology and cancer biology.

"There are colonies of pelagic tunicates which have taken shape like the finger of a glove. Each member of the colony is an individual animal, but the colony is another individual animal, not at all like the sum of its individuals... So a man of individualistic reason, if he must ask, 'Which is the animal, the colony or the individual?' must abandon his particular kind of reason and say, 'Why, it's two animals and they aren't alike in any more than the cells of my body are like me. I am much more than the sum of my cells, and, for all I know, they are much more than the division of me.'"¹

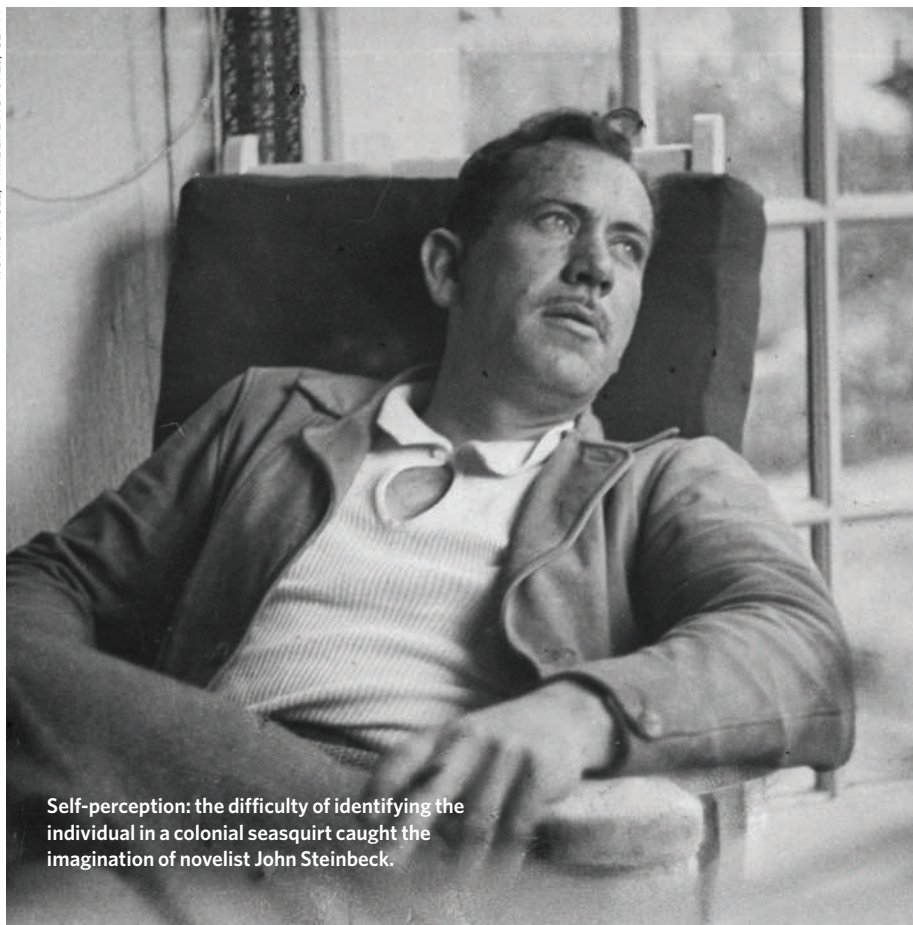
There is something wonderful, whether for a child or a great man of letters, in the sight of a rock pool brimming with strange and secret creatures. The pelagic tunicates that so enraptured John Steinbeck when he sailed the seas off California in 1940 with his friend, marine biologist Edward Ricketts, were particularly worthy of such wonder. Today, similar tunicates are providing scientific insight into the nature of the individual, in two seemingly separate realms. Both immunologists, interested in how the self is distinguished from non-self, and evolutionary biologists, trying to establish the level of organization on which natural selection acts, have turned to the tunicate *Botryllus schlosseri* for answers.

Botryllus schlosseri is a beautiful, cosmopolitan and enigmatic creature. Its life cycle starts with a tadpole-like animal that possesses a rod of elastic tissue called a notochord — the evolutionary precursor to the backbone. This tadpole picks a suitable patch of rock or seaweed frond on which to settle and, in a bizarre and seemingly atavistic transformation, metamorphoses into a polyp-like creature. The adult then begins to bud, producing genetically identical offspring called zooids. Once mature, the zooids produce their own buds. In a coordinated weekly orgy of death and regeneration, the old zooids perish and shrink back to nothing while the new take over.

The entire colony of zooids is connected by a system of blood vessels and is sheathed in a jelly-like tunic. Underneath this mantle, the filter-feeding zooids cluster prettily like petals around communal siphons, through which they expel water. The question of where the individuality lies in such a creature is not easy to answer.

Irving Weissman, of Stanford University, California, has been studying the *Botryllus* colonies of nearby Monterey Bay for 30 years. Now an eminent stem-cell biologist, he began his career with a keen interest in immunology. His interest drew him to the strange interactions that take place when the *Botryl-*

P. STACKPOLE/TIME LIFE PICTURES/GETTY



Self-perception: the difficulty of identifying the individual in a colonial seasquirt caught the imagination of novelist John Steinbeck.

lus colonies infringe on each other's borders

When the blood vessels of two colonies make contact, one of two things can happen. The two systems can fuse, producing a new colony that contains cells of two different genetic make ups — a chimaera. Or an inflammatory reaction can take place in which the interacting blood vessels are destroyed and a scar is formed that prevents fusion. Whether the colonies fuse or reject each other is determined by the genetic make up of each colony. Together these responses look, to an immunologist's eyes, uncannily like the mechanisms by which a body accepts or rejects an organ transplant.

In the 1980s, Weissman's team found that the fusion–rejection decision depended on a genetically encoded system that behaved very much like the major histocompatibility complex (MHC) in humans². The MHC, which displays specific protein fragments on the surface of cells, acts a bit like an identification tag and determines whether the tissues of a human organ donor are accepted or rejected by a recipient. But why do the colonies try to fuse in the first place?

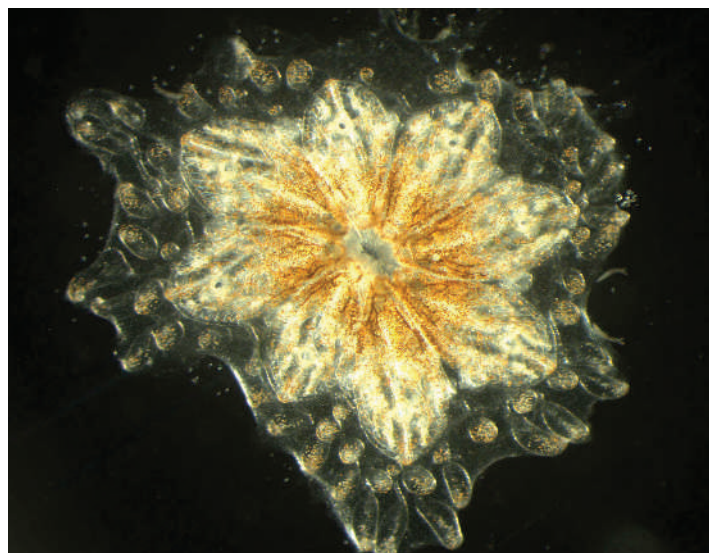
Skill sharing

One suggestion is that a combination of two (or more) tunicates is better than one. For example, in the lab, if a colony that thrives best at 15 °C merges with one that thrives at 25 °C, then the fused colony will do pretty well at both temperatures. “The chimaera is very flexible,” says Baruch Rinkevich, a biologist at the National Institute of Oceanography in Haifa, Israel. Pooling talents in this way would make the colonies better able to deal with environmental change. It also means that a colony can spread itself over a larger area — handy if you get nibbled by hungry fish. But how can colonies merge their characteristics?

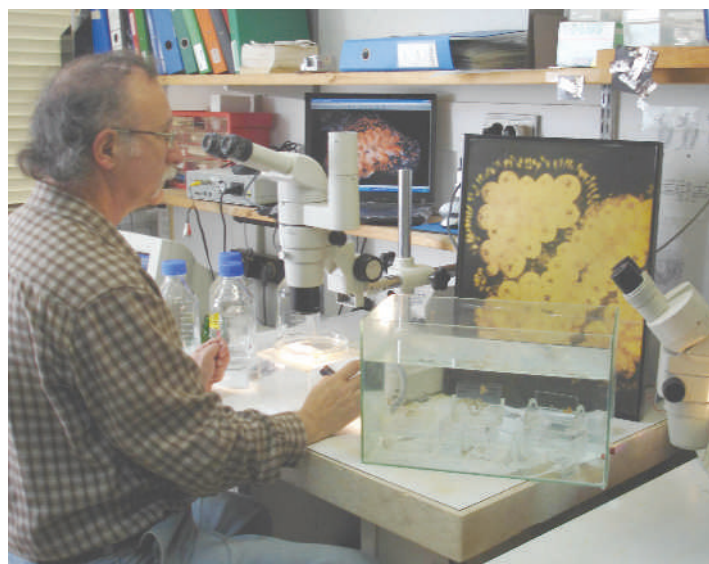
The answer lies in the tunicate's extraordinary ability to regenerate. The colonies can regrow themselves every seven days thanks to a particular sort of stem cell. In adult vertebrates, stem cells divide to produce both more stem cells and cells of various different types specific to a particular tissue; skin stem cells can produce various sorts of skin cell, for example. But take a tiny scrap of blood vessel from one of these tunicates, and it will regenerate the entire animal. In other words, some of the tunicate's adult stem cells behave like our embryonic stem cells; they have the power to make any cell type in the tunicate 'body'.

In creatures that develop once, such as humans, such powerful stem cells are needed only early on in development. But in a creature that regrows itself on a weekly basis, they must be on call all the time. “It's totally backwards,” says Anthony De Tomaso, a former postdoc of

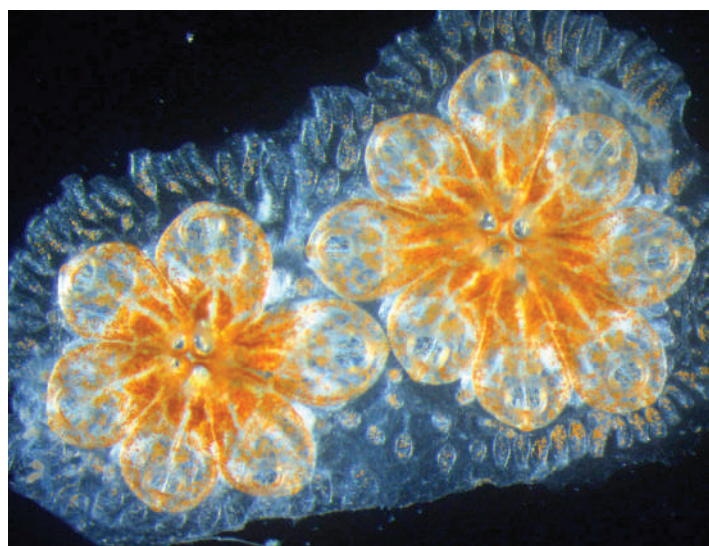
Baruch Rinkevich (middle) sees parallels between cancer and the behaviour of stem cells of the tunicate *Botryllus schlosseri* (top and bottom), which run riot in the tissues of a closely related species.



G. PAZ



G. PAZ



G. PAZ

Weissman's who has taken up the reins of the Stanford tunicate lab. "But it sets up this situation where you can have development occurring all the time."

When two colonies fuse, the stem cells of each are free to build tissues in what used to be a separate colony. This mixes up the different cells and gives the 'new' tunicate the benefit of a greater range of gene variants without all the fuss and bother of sex. But the fusion raises questions about what level of organization natural selection now acts on — the genetically distinct cell lines originating from the two colonies, or the merged colony as a whole?

It was another Weismann, differently spelled, who a century earlier argued that natural selection does not take place within the body. The nineteenth-century German biologist August Weismann divided the cells of multicellular creatures into two types: 'somatic cells', which make up almost all the parts of the body, and the cells of the 'germ line' — a small minority that produce just egg and sperm.

Crucially, he said, there is a wall — the Weismann barrier — between the soma and the germ line. This barrier ensures that somatic cells can never contribute to the next generation, and prevents any of the creature's acquired characteristics from being passed on to its descendants. Weismann's work was fundamental to the rediscovery of mendelian genetics and to twentieth-century understanding of darwinian evolution. This viewed natural selection as acting on the individual organism, with the quality of the creature's germ line tested by the fitness of its genetically identical somatic cells.

The idea of the individual as the only unit of selection has since been challenged by biologists advocating the role of higher and lower

levels of organization. Group selectionists hold that evolution can select for attributes that benefit a group or species as a whole. Meanwhile, gene selectionists argue that evolution can be best understood in terms of choices between 'selfish' genes competing to copy themselves. The tale of the tunicates suggests another level, where selection acts directly on cell lines.

One champion of this idea is Leo Buss, a biologist at Yale University in New Haven, Connecticut, whose decades of work on the topic have only recently begun to be addressed by molecular biologists. Buss argues that in protists, fungi, plants and 19 of the 33 different animal phyla around today, the Weismann barrier can leak³. In these organisms, somatic cells can become germ cells, and thus genetic changes acquired during development within the soma can become heritable. Natural selection can thus act on competing cell lineages within the body, favouring those that get into the germ line.

One-way street

This competition began more than half a billion years ago, when a group of single-celled organisms made the great transition to forming a multicellular body. Understanding how natural selection resolved the competitive conflicts between cells, and between cells and the individual creature, says Buss, is key to understanding how multicellular organisms evolved into the forms we see today.

With Buss's work in mind, Weissman's team at Stanford traced the destiny of cells in fused tunicate colonies. The group found that, in some cases, cells from one colony seemed to completely replace the body tissues of another⁴.



In others, cells from one colony sneaked into the gonads of the other, replacing the host's germ line^{5,6}. In this fate worse than death, the tunicate is not just stopped from propagating its own genetic line, but effectively forced to churn out a competitor's offspring.

It is this dreadful spectre that gives rise to the tunicate's rejection reaction. To avoid being parasitized, *Botryllus* colonies have acquired a well-developed 'sense of self' — the genetically encoded tissue-matching system, now called the FuHC system, that Weissman's team found in the 1980s. One colony will only permit another to fuse with it if its FuHC tissue-matching genes are sufficiently similar. The FuHC gene, like the MHC genes, comes in thousands

D. KRAFT

THE BIGGEST PARASITE IN THE WORLD?

In 1876, a Russian vet called Nowinsky performed some rather disgusting experiments. He was interested in a strange cancer now called canine transmissible venereal sarcoma, which affects the faces and genitalia of dogs.

By transplanting tumours from the genitals of one dog to another, Nowinsky confirmed the initially surprising suggestion that this tumour was transmissible. Later, other scientists showed that the cancer was transmitted by its own cells, which have the ability to detach themselves from the

tumour they start off in, in order to emigrate to a new host.

The disease probably originated from a single tumour in a single dog⁹. That dog is long dead, but the cell lineage it fostered is alive and well in the genitals of countless dogs and foxes all over the planet. It is common in strays in Japan, the United States, Europe, China, the Far East, the Middle East and parts of Africa. In some places it infects up to a third of the animals that might be susceptible. It may well be the biggest parasite in the world.

In most cases, the dog's

immune system eradicates the cancer in months. But in puppies and dogs whose immune systems are depressed, the cancer spreads throughout the body — just like the *Botryllus* super-predator stem cells. Problems with the immune system might also underlie another transmissible cancer — one that is devastating populations of the Tasmanian devil in Australia. Devil facial-tumour disease spreads when the devils bite each other during fights or courtship. Like the dog tumour, it probably originated in the body of a single individual¹⁰; unlike the canine cancer it kills

almost every devil it infects. This may be because the devil population is so inbred that catching the cancer is like getting a tissue transplant from a close relative; the cell lineage can escape the immune system's radar.

Humans, too, can occasionally catch a deadly cancer — if they receive cancerous organ transplants and their immune systems are subverted by immunosuppressive drugs. For this reason, organs are normally no longer harvested from the bodies of people who are suffering from or have died from cancer.

C.A.



Irving Weissman (left) and Tony De Tomaso (right) have shown that a tissue-matching system analogous to that in humans operates in tunicates.

of different versions; two colonies are only likely to be a match if they are closely related. In this case, the costs of germline takeover are much diminished. If incoming stem cells do hijack the germ line, they will be closely enough related for it not to matter too much; most of the host's genes will effectively still get through.

De Tomaso and colleagues think that the FuHC operates in much the same way as the human MHC does in natural killer cells. Unlike T cells — white blood cells that actively look for foreign proteins associated with the MHC — natural killer cells look for 'missing self'; that is, for cells that lack the normal MHC. Although the molecular players appear to be different in the tunicate, the same logic seems to apply when one colony merges with another.

All creatures that can live as chimaeras, including fungi, flowering plants and primitive animals such as sponges, have some kind of ability to recognize foreign cells. Luis Cadavid, a biologist at the University of New Mexico in Albuquerque who has studied another colonial sea creature, a hydrozoan called *Hydractinia* finds that it, too, has a set of markers for selfhood. The commonality of this trait does not mean, however, that it has a shared origin. Although FuHC may function like the MHC, the sequence data of de Tomaso's team's shows that the two systems are not related⁷.

"My impression is that there is a common theme to recognizing self versus non-self, but the mechanisms may have different origins," says Cadavid. Louis du Pasquier, a biologist at

the University of Basel in Switzerland who works on the evolution of adaptive immunity, agrees. What the evidence points to, says Du Pasquier, is a common selection pressure — originally for a way to maintain the integrity of the self in the face of invasion, not by bacteria and viruses, but by members of your own species.

Self aware

This applies to humans too, argue De Tomaso and Weissman. Humans can play host to competing cell lines in a number of ways. For example, in an organ transplant or blood transfusion, blood stem cells from the donor can quickly colonize the recipient's body and may hang around indefinitely. But there are natural ways too. Although not an example of chimaerism, cancer involves a particular line of somatic cells declaring an independent identity from the body around it. In order to avoid alerting the adaptive immune system to their aberrant nature, some cancer cells stop expressing MHC genes.

Rinkevich and his team have studied, as an analogous process, cells from *Botryllus* running riot in the bodies of a closely related species of tunicate called *Botrylloides*. "It's like cancer," says Rinkevich. "It does have parallels, no doubt about it," agrees De Tomaso. "Cancer is somatic cell parasitism. It is selection inside a body for a variant."

Weissman thinks that *Botryllus* has a lot to teach us. Cancer researchers now think that our bodies' stem cells play a key role in both the origin and development of cancerous tumours⁸, and Weissman sees intriguing parallels between such stem cells and the most highly successful of *Botryllus*'s 'winner' stem cells. He is keen to uncover the genes that give these cells their super-predatory powers. "I wouldn't be surprised if, when we finally isolate these genes, related genes are found in pathways that allow cancer cells to develop," he says.

Cancer is not generally a route to true immortality, because it is normally limited to a single body with a finite lifespan (although not always, see "The biggest parasite in the world?"). To persist beyond this, stem cells need to get out of the soma and into a germ line — any germ line. In theory, the easiest way for this to happen would be through pregnancy. Scientists have known for years that in many mammals, stem cells can cross from fetus to fetus in multiple pregnancies. Stem cells can also break out from fetuses into the mother's blood. Most are hunted down and destroyed by the mother's immune system, but usually some remain — although no one knows why or how. A woman

who has had several children could be a mixture of many different cell lineages.

Such embryonic stem cells could in principle colonize the germ lines of subsequent embryos, or even, conceivably, the mother's germ line. But germline chimaerism seems not to happen. And to Weissman, that implies that our MHC-based systems of selfhood, like those of the tunicates, know how to stop such takeovers. "The fact that we don't all walk around as germline chimaeras would imply that there is a protection mechanism" says Weissman.

De Tomaso suggests that the need to stop these cells from taking over could perhaps explain a long-standing mystery in immunology. The diversity seen in MHC genes is much greater than can be explained by the needs of the adaptive immune system. Perhaps the threat of stem-cell parasitism is to blame? So far, de Tomaso admits, the idea is just speculation, and there are a number of other theories that attempt to explain MHC diversity. But he's still intrigued by the idea that the need to distinguish ourselves might be a key driver of natural selection at the level of the organism, pushing people towards uniqueness. "Everything I see," he says, "suggests that there is selection for you to be as rare as possible."

And so the maintenance, even exaggeration, of individuality in humans and the circumstances under which other creatures will surrender themselves to greater groups may be underpinned by common mechanisms. The levels of the self, the cell and the gene all interact under rules that were laid down half a billion years ago, and yet are invigilated round the clock today. Understanding the drives of the

"Everything I see suggests that there is selection for you to be as rare as possible."

— Anthony De Tomaso

self, the cell and the gene, and recognizing that cells are not just inert building blocks but entities with at least a capacity for

selfdom, is, as Steinbeck wrote when pondering his tunicates, "the basis for a far deeper understanding of us and the world".

Claire Ainsworth is a senior reporter for *Nature*.

1. Steinbeck, J. *The Log from the Sea of Cortez* (Penguin, 2000).
2. Scofield, V. L., Schlumpberger, J. M., West, L. A. & Weissman, I. L. *Nature* **295**, 499–502 (1982).
3. Buss, L. W. *Proc. Natl Acad. Sci. USA* **80**, 1387–1391 (1983).
4. Stoner, D. S. & Weissman, I. L. *Proc. Natl. Acad. Sci. USA* **93**, 15254–15259 (1996).
5. Stoner, D. S., Rinkevich, B., Weissman, I. L. *Proc. Natl Acad. Sci. USA* **96**, 9148–9153 (1999).
6. Laird, D. J., De Tomaso, A. W. & Weissman, I. L. *Cell* **123**, 1351–1360 (2005).
6. De Tomaso, A. W. et al. *Nature* **438**, 454–459 (2005).
7. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. *Nature* **414**, 105–111 (2001).
8. Katzir et al. *Oncogene* **1**, 445–448 (1987).
10. Pearse, A.-M. & Swift, K. *Nature* **439**, 549 (2006).

A SINKING FEELING

The floods are getting worse in Tuvalu. As scientists argue over climate change and struggle to measure rising seas, **Samir S. Patel** meets the locals of this tiny island nation.

On a Friday afternoon in late January, the phone rings in the Filamona Guest House, one of the few places to stay in the Pacific archipelago of Tuvalu. Hilia Vavae, the director of the Meteorological Office, sounds excited in her modest, laid-back way. "There is a..." she struggles for the right word "...flooding!"

The floods on Funafuti, the main atoll of Tuvalu, are caused by 'king tides' — the highest high tides of the year. They percolate up through the porous limestone, soaking the islets from the inside out. And they draw journalists, scientists and environmental advocates from around the world, all keen to see what is happening on the front line of climate change. If global sea levels rise as predicted, the 11,000 Tuvaluans will be among the first to see things go awry. Yet even here, attitudes and responses to the impending catastrophe vary in complex and sometimes surprising ways.

Next to the guest house is the country's sole airstrip, built by the US military in the Second World War. Being the only open space on Funafuti, the airstrip provides a breezy place to sleep on muggy nights, not to mention a pitch for football and *te ano*, a volleyball-like game.

Following Vavae's directions, I walk across the airstrip and, rolling up my trousers, on into a clear tidal lake. Two kids run over to sail toy boats made of green Victoria Bitter beer cans pounded flat. The salt water creeps steadily along the dirt track next to the runway and surrounds the nearby buildings, including the diesel power station, a handful of pungent concrete pigpens, and the office where Vavae and her staff collect data on sea level and beam them to the Australian Bureau of Meteorology's National Tidal Centre (NTC), thousands of kilometres away in Adelaide.

In January and February, Tuvalu experienced some of the highest tides ever recorded there, nearly 1.5 metres above mean sea level. They were caused in part by the convergence of natural short- and long-term tidal cycles, but were boosted, perhaps, by the effects of global warming. Although the tides are not especially high compared with Newfoundland's Bay of Fundy or even the English Channel, they are pretty alarming in a country where the highest ground is just 5 metres above sea level and most is much less than that.

A mean sea-level rise in Tuvalu of just 20 to 40 cm in the next hundred years would signif-

icantly increase the frequency and depth of saltwater flooding and accelerate coastal erosion. It would threaten the Tuvaluans' food and housing, poisoning the pits where they grow giant swamp taro plants and undermining buildings. It could make the country simply uninhabitable.

Rising damp

There are two tide gauges in Tuvalu. One, operated by the University of Hawaii until 1999, sits on a small concrete wharf behind the three-storey Taiwanese-built government building. In 1993, the NTC installed a more modern and accurate gauge a few kilometres north at the country's only deepwater wharf. One of twelve in the South Pacific, this gauge should in theory provide quantitative confirmation that Tuvalu is being engulfed, as the king tides and the wet cuffs of my trousers suggest.

But in 2000 an NTC analysis reported a negligible increase of 0.07 mm a year over the past two decades from the University of Hawaii gauge, and a drop in sea level from the seven years of NTC data¹. It was clear that the El Niño/Southern Oscillation (ENSO), which drives down sea level in the western Pacific, affected both of these records. And the international environmental group Greenpeace asked John Hunter, a climatologist at the University of Tasmania, to have another look at the data. When he adjusted for ENSO and the vertical movement of the Hawaii gauge, which is thought to be sinking, Hunter found a sea-level rise of around 1.2 mm a year².

Water, water everywhere: Tuvaluans think the floods are getting worse, but it is hard to measure changes in sea level accurately.





M. PALEY/CORBIS; SATELLITE IMAGE BY GEOEYE

Shore thing: the snaky islet of Fogafale is Tuvalu's capital and hosts the country's airstrip (inset).

Hunter's figure is consistent with the global estimate of the Intergovernmental Panel on Climate Change (IPCC): 1 to 2 mm a year for the twentieth century³. But the Tuvalu estimates are based on a couple of gauges and a reasonably short record, points out John Church of the Commonwealth Scientific and Industrial Research Organization (CSIRO) in Hobart, Tasmania, who was one of the lead authors of the chapter on sea level in the IPCC's most recent assessment.

Recently, Church and his CSIRO colleague Neil White have moved to a more regional approach. They have combined records from tide gauges around the world, some of which date back as far as 1870, with satellite altimeter data to assess regional variation in sea-level rise. Their results for the South Pacific are in line with the Hunter and IPCC estimates⁴, and they are now looking specifically at Tuvalu and other small island nations.

"The thing that really interests me is how you reconcile the relatively low estimates of sea-level rise, which are the same order as what's happening in the rest of the world, with the anecdotal observations from Tuvalu," says Hunter. "It seems that the flooding reported there is bigger than 2 mm a year. The extremes of high tide could be getting bigger relative to the mean sea level, although that's disputed at the moment."

Phil Woodworth of the Proudman Oceanographic Laboratory in Liverpool, UK, has studied high tides in the Pacific to see whether they have changed since the 1970s. He says the peaks seem to be increasing, but by no more

than mean sea level⁵. "The extremes at Tuvalu seem determined primarily by the slower changes in the oceans due to ENSO and the long-term changes in mean level," he says.

The disagreements are unlikely to be resolved soon. "We're going to be waiting around for a fair while before our estimates of sea-level rise become statistically meaningful," says Bill Mitchell, manager of the NTC. "Everyone presses you to give a number. We put a vast amount of effort into telling people that you should not be using numbers yet."

Treading water

Vavae has a similar opinion of the numbers, but perhaps for different reasons. "I always tell my people that it's not the data that you look at. You have to actually rely on your eyes," she says, glancing outside to where the salt water is once again creeping across the yard. "A person who experiences it has got a much better feel or much better knowledge of what is happening than someone who is 100 miles away."

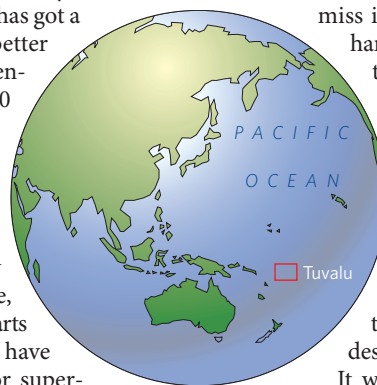
When Vavae started working for the Meteorological Office in 1981, the saltwater flooding was no worse than that from heavy rain. Now it is extensive, regularly inundating large parts of the island. The floods have almost reached the *fusi*, or super-market; seaweed is starting to appear in the more commonly flooded areas, Vavae tells me.

The Tuvaluans have mixed reactions to these changes, and the scrutiny that comes with them. A few days after witnessing the king-tide

flood by the airstrip, I talked to Carol Farbotko on the verandah of the Filamona. Farbotko, a cultural geographer from the University of Tasmania, is working on a thesis on cultural responses to climate change on the islands. In her interviews with officials and community leaders, she has found that climate is a vague, long-term concern. People get much more worked up about problems such as waste disposal, a fetid and ubiquitous problem; overpopulation; and the accelerating erosion of traditional culture in the age of the Internet and DVDs. Even frequent workshops on climate and the dangers of accelerating sea-level rise fail to provoke a sense of urgency.

"Oh, it's a very important concern," is the standard and slightly mechanical response to questions on climate change, Farbotko says. In her experience, and in mine, some Tuvaluans refuse even to talk about climate, or dismiss it with a weary wave of the hand. Tuvalu is a deeply Christian country, and some islanders put their faith in the promise God made to Noah in Genesis 9:11: "And I will establish my covenant with you; neither shall all flesh be cut off any more by the waters of a flood; neither shall there any more be a flood to destroy the Earth."

It would not take a very large breach in that covenant to wash Tuvalu away. Fogafale, the largest of the islets that make up the Funafuti atoll, is an elegant snake of land about 10 kilometres long. Even on a borrowed bicycle with a hard seat and half-flat





S. S. PATEL

Where next? If Tuvalu is flooded, its citizens could relocate, but their culture might be lost.

tyres it is easy to get from one end to the other in a morning. The ride is comfortable, because the roads were resurfaced with a windfall from the sale of Tuvalu's appealing Internet country code, '.tv', although the speeding motorcycles and cars no longer have to slow for potholes.

During high tides, waves crash from the lagoon on to stretches of the new tarmac, spreading trash, coral rubble and other detritus across it. There was a time, I'm told, when the road behind the government building was set back a few metres from the lagoon, but now it is right on the sea. In itself this says nothing about sea level, because the islets are constantly changing shape: a spit forming here, an island tip disappearing there. Arthur Webb, a coastal ecologist in the Fiji offices of the South Pacific Applied Geoscience Commission, makes the point to me graphically by flipping through aerial and later satellite photographs from the 1940s on. "These really are quite natural processes," he says. "It's part of living on a soft-shored island."

Swept away

As Webb shows me images with Second World War seaplanes and torpedo boats sitting in the lagoon, he explains how, in 1943, US troops built a makeshift seawall and land-reclamation project along the length of the lagoon fore-shore. After the troops left, locals built homes and roads on this 25- to 30-metre-wide stretch of rubble, which then began to erode. Jetties and channels further altered coastal erosion and deposition patterns.

The move from portable, short-lived thatch houses to Western-style concrete-block homes has added to the difficulties. Despite regulations, the shores of Tuvalu have been ransacked for aggregate for construction — an activity that makes the prospect of flooding far worse. "Beach mining is a disaster," says Webb.

Those sceptical about Tuvalu's plight, including amateur scientist Willis Eschenbach, seize on local explanations such as mining to assert that fears about sea-level rise are created by hysterical journalists and environmental groups looking for a *cause célèbre*. Eschenbach, who carried on a spirited debate with Hunter in the journal *Energy and Environment*, has concluded that sea-level rise in Tuvalu is an

illusion. He has used that conclusion to support an argument that there is no clear evidence for climate change.

Webb, on the other hand, thinks that poor coastal management and climate change are acting in concert. Oddly, when he presented this sinister synergy to the *falekaupule*, or council of elders, in Fogaale, their reaction was something like relief; a remote and difficult problem became a local and understandable one. They learned there were things they could do to mitigate erosion — they could further regulate beach mining and carefully dredge the lagoon for the island's aggregate needs, for example.

Many scientists and people within the government insist that Tuvalu must step up the implementation of 'no-regrets' policies — activities that make sense whether the seas rise or not. These include introducing salt-tolerant crops and dealing with the island's highly visible trash problem. The process is painfully slow.

Distressed that its capacity for action was limited to its own shores, in 2000 Tuvalu joined the United Nations for the express purpose of highlighting climate change. It was a considerable expense for a country that had a GDP at the time of just US\$12.2 million. But membership allows Tuvalu to play a role as the most vocal and insistent of the small island developing nations, positioning itself as the conscience — or pest — of climate negotiations. "The current strategy is to continue making noises in the international forums," says Prime Minister Maatia Toafa.

Exit strategy

In addition to championing increased adoption of the Kyoto Protocol, which aims to curb greenhouse-gas emissions, Tuvalu also wants to discuss immigration policies with Australia and New Zealand. More open policies would provide both economic opportunities and the possibility of a new home if, or when, the islands become uninhabitable. Critics charge Tuvalu with using the sympathy generated by its position to increase the number of Tuvaluans living abroad and the remittances they

send home. Toafa, a casual and jovial fellow who kicks off his sandals and props his feet on the table as we speak, is candid about their intentions. "It will work both ways," he says. "One, as an opportunity for people to go and develop their lives there, and secondly as a way of easing this resettlement problem."

Ideally, Toafa would like to buy land in New Zealand or Fiji to resettle the entire nation. "Because we love the sea, we need a place close to the sea. And we know these are very expensive places," Toafa says. But relocation is more than a logistical and economic problem. It threatens their national and cultural identity — "unless we can develop an underwater Tuvalu," he adds with a high-pitched laugh.

So far, New Zealand has offered to accept more Tuvaluans, but that brings the total to just 75 a year, and even then it is as part of a labour programme, not resettlement. Despite reports in the press that the evacuation of Tuvalu is already under way, very few are going. "I don't know if anyone wants to leave," says PePETua Latasi, climate-change coordinator for the environment department. "People are saying Tuvalu will be gone in 50 years' time, but I doubt it." Still, even Latasi, who is optimistic about both local efforts and international mitigation, concedes that if the expected rises in sea level coincide with increases in cyclone activity, the prospects are bleak.

On another day of king tides, I walk down to the end of the airstrip, where floods overflow the edge of a trash-filled pit. Three boys are trying to turn over a half-submerged rowing boat. It capsizes, but they paddle back and have another go.

Samir S. Patel is a freelance science journalist and photographer based in New York City.

"I always tell my people you have to rely on your eyes."
—Hilia Vavae

1. Mitchell, W., Chittleborough, J., Ronai, B. & Lennon, G. W. *Q. Newslett. S. Pacific Sea Level Clim. Monit. Proj.* **5**, 10–19 (National Tidal Facility, Australia, 2000).
2. Hunter, J. *A Note on Relative Sea Level Change at Funafuti, Tuvalu* (2002); available at <http://staff.aecrc.org.au/~johunter/tuvalu.pdf>.
3. Intergovernmental Panel on Climate Change *Climate Change 2001: The Scientific Basis* Ch. 11 (2001).
4. Church, J. A. et al. *J. Clim.* **17**, 2609–2625 (2004).
5. Woodworth, P. L. & Blackman, D. L. *J. Clim.* **17**, 1190–1197 (2004).

BUSINESS

Biologist aims to ease the pain for entrepreneurs

Biotech start-ups face a struggle to survive. Virginia Gewin reports on an initiative that brings cash and experience together to improve their chances.

The road from invention to commercially successful company can be a rocky one — and biologist Leroy Hood has travelled it many times.

Inventor of the automated DNA sequencer, Hood has helped to found 13 companies — including one of the world's top biotechnology firms, Amgen. His knowledge of the potential pitfalls faced by scientists making the transition from lab to commerce has led him to set up Accelerator Corporation, a venture that Hood believes offers a fresh approach to technology transfer.

"The idea is to accelerate early-stage discoveries and push successful ones on to the next stage of funding," says Hood.

Accelerator is affiliated to the Institute for Systems Biology (ISB) in Seattle, Washington, which Hood co-founded in 2000. The company combines the institute's scientific acumen with six venture-capital partners to provide support and advice for new biotech businesses.

The initiative focuses on a perilous time in a company's evolution: the very early stages of development, when academics with a promising idea need money and management expertise to help the firm take shape.

Enticing venture capitalists to invest in emerging biotech firms is currently problematic. Investors are still smarting from losses made earlier this decade on over-hyped start-up firms, and so are inclined to limit their risks by backing more mature companies. And many venture-capital funds are too large to bother spending valuable time cultivating very small firms.

As a result, funding in the United States for early-stage biotech companies last year accounted for just 2% of the \$1.7 billion that venture capitalists invested in the sector, according to research firm VentureOne of San Francisco. In that climate, even the best ideas can struggle to win financial backing.

That has left many US academic institutions actively seeking a way to secure backing for their ideas. More than 90 of them have established 'gap funds' to help their start-ups get under way — and many of the institutions call on venture capitalists for advice and

support when they deploy these funds.

Accelerator takes this idea a stage further by getting the venture capitalists to back the firms directly. It was launched in 2003 by Hood and one of his former graduate students, Michael Steinmetz, who worked for ten years as a research manager at drug firm Roche and is now a partner at MPM Capital in San Francisco. To date, Accelerator has drawn in \$22 million of investment capital from six venture-capital groups, including MPM.

As well as some seed funding, the companies taken on by Accelerator get advice and support from the investors on how to position themselves to win more substantial backing later on. The ISB, meanwhile, helps to steer the companies scientifically in exchange for some equity, which Hood hopes will end up sprouting healthy returns for the institute's endowment.

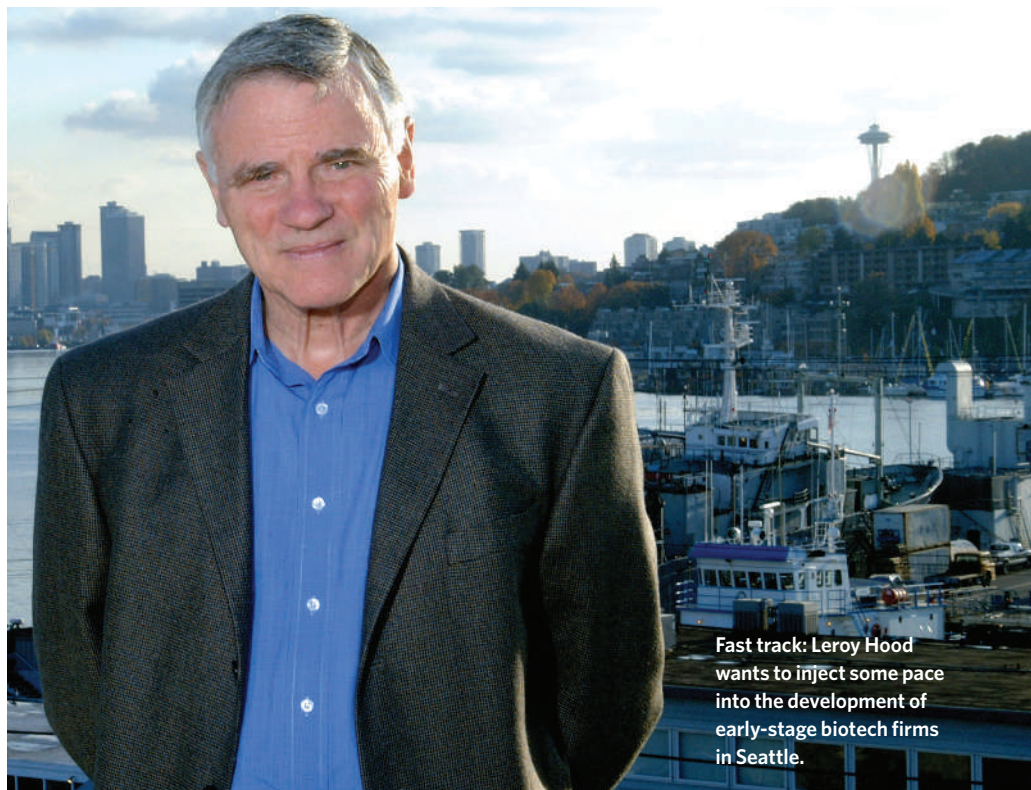
Carl Weissman, a seasoned biotechnology

entrepreneur and another partner at MPM Capital, directs Accelerator's small core staff. His team helps to run the nascent companies while their scientists focus on research. Other 'incubator' arrangements of this type often "provide nothing above rudimentary management services, such as answering the phone", Weissman claims. But drawing on their venture partners' expertise, Accelerator staff recruit managers, negotiate licences, write business strategies, hire lawyers and even help to arrange future rounds of financing.

One of Accelerator's first clients was Seattle-based VLST, co-founded by biochemist Craig Smith. The company is using research on virus survival to develop more streamlined techniques to find drugs for both autoimmune and inflammatory diseases. For an undisclosed stake in the company, Accelerator has helped Smith to hire a chief executive and a chief financial officer.

So far, Accelerator has taken on five companies, two of which are in the process of securing their first round of outside venture funding. One of these will be among the largest financings in the entire region this year, says Weissman — significantly higher than the \$10 million to \$20 million typically raised for biotech firms at this stage in their development.

The Accelerator model gets venture-capital groups more closely involved with early-stage companies than is usually the case. At the



Fast track: Leroy Hood wants to inject some pace into the development of early-stage biotech firms in Seattle.

G. NYSTROM/ISB

Massachusetts Institute of Technology (MIT), for example, investors don't put in funds but do offer advice to managers of its gap fund, which is run from the Deshpande Center. Set up in 2002 with a \$20-million gift from optics entrepreneur, Desh Deshpande, this centre has so far given grants to about 50 projects. These have yielded nine companies, which between them have attracted \$40 million in venture funding.

But many universities have had a tougher time with their initiatives. When the Bio-Generator was founded in 2004 in St Louis, Missouri, for example, it said it hoped to be managing seven companies within two years.

"New ideas need new organizational structures."

So far it has found only three, and is being restructured to provide more intensive support for its charges — including more money to take their research forward.

"We now understand the importance of putting money into translational research while technology is still at the university, before making a commitment to create a company," says Marcia Mellitz, a BioGenerator board member who also runs the Center for Emerging Technologies, another St Louis business incubator.

Indeed, venture-capital groups can shy away from university-based projects that incubate new companies. Investors sometimes claim that the universities — which are keen to please academics and satisfy local political demands to create jobs — don't assess commercial potential objectively enough, and will continue to support ideas that don't meet expectations.

"A venture capitalist putting a company together is a whole lot different from what happens in a university setting," says Patricia Beckmann, chief scientific officer of Homestead Clinical Corporation, a blood-diagnostics company that originated at the ISB and is being backed by Accelerator. "Our feet are in the fire for returns," she says, referring to the demands that are placed on her company to meet milestones.

Institutions in other cities, from San Diego to St Louis, have visited Seattle to spy on how Accelerator is shaping up. "We're not saying it's the only model, but it's a model that works," says David Schubert, Accelerator's business officer. He acknowledges that it will take another five years to measure its success or failure.

Hood ascribes the idea to his frustration in the past with the amount of time it can take to obtain venture-capital funding. After three decades in biotechnology, he says he has learned at least one thing: "New ideas need new organizational structures." ■

IN BRIEF

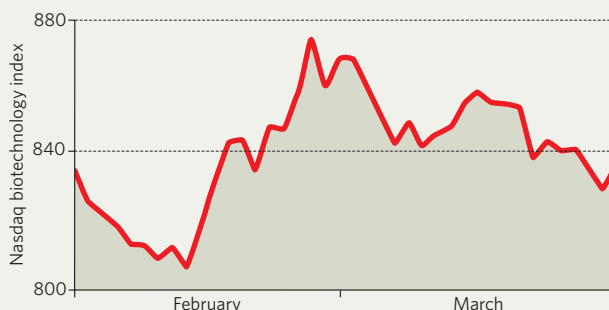
TOTALLY WIRED The United States has reclaimed pole position in a league table of the 'most-wired' nations, according to the World Economic Forum. Singapore, Denmark, Iceland and Finland fleetingly supplanted it last year as the nations where telecommunications and information technology had the widest and deepest reach, the forum's annual survey finds. It describes the United States as a 'powerhouse' in information technology that continues to set the standard for other nations.

GENERIC GRAB Europe's largest pharmaceutical company has carved out a piece of the burgeoning market for generic drugs, buying one-quarter of a fast-growing East European generics developer. Sanofi-Aventis has bought 24.9% of Zentiva, a Czech generics maker with 4,200 employees, for €430 million (US\$520 million). The move makes Sanofi-Aventis the largest single shareholder in Zentiva, which markets 270 drugs in eastern and central Europe, and has additional production sites in Slovakia and Romania.

NUCLEAR SELL-OFF The UK government has confirmed that it will sell its nuclear clean-up business, British Nuclear Group, by autumn 2007. The group is currently part of the state-owned BNFL. The government also upped its estimate of the total future cost of cleaning up existing UK nuclear sites from £56 billion (US\$98 billion) to £70 billion — the British Nuclear Group is likely to be well-positioned to win contracts under this huge clean-up programme

MARKET WATCH

BIOTECHNOLOGY STOCKS



Wood Mackenzie, an Edinburgh-based research and consulting firm, reviews recent trends in biotechnology stocks.

The Nasdaq biotechnology index rose in February and dropped back in March, but the overall trend remains positive, with a gain of 6% since the start of the year. February saw strong 2005 financial results posted by many of the major biotech companies, bolstering investor confidence in the sector.

The changing fortunes of Biogen Idec in Cambridge, Massachusetts, had a particularly strong influence on the index. The company's year so far has been a tale of two antibodies. Its shares rose 12% in February when the US Food and Drug Administration (FDA) approved its anticancer antibody Rituxan to treat rheumatoid arthritis and an FDA advisory committee recommended that the antibody Tysabri should return to market to treat multiple sclerosis. Tysabri had been withdrawn from sale in February

2005 on safety grounds. The FDA subsequently delayed making a final decision on the drug, causing Biogen's shares to shed much of their earlier advance.

Other share movements were dictated by the results of clinical trials. Adolor of Exton, Pennsylvania, saw its share value rise by more than half after good trial results for its post-surgery bowel treatment Entereg. Another Pennsylvania company, Novavax in Malvern, enjoyed a whopping 79% increase in its share price after presenting promising preclinical data on its early-stage bird-flu vaccine.

But Antigenics of New York lost almost half its share value when experimental cancer drug Oncophage failed to meet expectations in kidney cancer patients. And unexpected side effects from hepatitis C treatment valopicitabine shaved nearly 40% from the share price of Idenix Pharmaceuticals of Cambridge, Massachusetts. ■

Avoiding hazards of best-guess climate scenarios

SIR — Your Special Report “The costs of global warming” (*Nature* **439**, 374–375; 2006) gives an unbalanced picture of the emissions scenarios developed by the Intergovernmental Panel on Climate Change (IPCC).

In contrast to the claim that these scenarios are outdated, a recent peer-reviewed assessment has concluded that, with a few notable exceptions, they compare reasonably well to recent data and projections for gross domestic product, population and emissions (D.v.V. and B.O’N. *Clim. Change*, in the press. doi: 10.1007/s10584-005-9031-0; see www.iiasa.ac.at/Research/PCC/pubs/vanVuuren&ONeill2006_CC_uncorproof.pdf).

Although we believe that progress on scenario development can and will be made, the elements of ‘up-to-date’ economic theory identified as overlooked — “how future societies will operate, how fast the population will grow, and how technological progress will change things” — are either too vague to be meaningful, or are issues the community has been dealing with for years. The Energy Modeling Forum has a 30-year history of model comparisons, exploring the implications for climate policy of a range of rates of economic growth and technological change (D. W. Gaskins and J. P. Weyant *Am. Econ. Rev.* **83**, 318–323; 1993, and J. P. Weyant *Energy Econ.* **26**, 501–515; 2004).

It is not correct to imply that the scenarios only use market exchange rates, or that they all assume that “the economies of poor countries will quickly catch up with those of rich nations”. Some scenarios are also reported in terms of purchasing-power parity exchange rates in the original 2000 IPCC Special Report. The debate on the emissions impacts of alternative exchange rates in economic modelling is not conclusive, but such impacts are likely to be small compared with the influence of technology, lifestyle and climate policies. And in no scenario do developing countries become as affluent as industrialized ones.

The assumed degree of catching up in the scenarios covers a wide range of possibilities. Focusing on a small number of most-likely futures ignores lessons from history: if the world always worked according to best-guess projections, we would now be living with nuclear power too cheap to meter and no ozone hole.

Arnulf Grubler*†, **Brian O’Neill***‡, **Detlef van Vuuren**§

*International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria

†School of Forestry & Environmental Studies, Yale University, New Haven, Connecticut 06511, USA

‡Watson Institute for International Studies, Brown University, Providence, Rhode Island 02912, USA

§Netherlands Environmental Assessment Agency, PO Box 303, 3720 BA Bilthoven, The Netherlands

Physician-scientists are needed now more than ever

SIR — Your News Feature “Them and us no longer” (*Nature* **439**, 779–780; 2006), about the narrowing divide between medical doctors and PhD scientists, ends with a disturbing assumption: that “the era of the physician-scientist [is drawing] to a close”. As medical doctors who have spent our careers in science, we strongly favour programmes that expose PhD researchers to the realities of clinical medicine. But these efforts do not eliminate the need for the classic physician-scientist, who obtains hands-on training in medicine as well as an advanced science education.

The recent rapid decline in the number of physician-scientist trainees in the United States and elsewhere is a serious problem that has been largely ignored. Now more than ever, we need them to continue bridging the intellectual and conceptual gap between medical doctors, seeking to understand the potential for science to deliver better care, and PhD researchers with an increasing interest in the same goals.

The disturbing trend in which fewer novel therapeutics are reaching the clinic will also not be halted without a robust pipeline of physician-scientists. Unless their training is restored, everyone stands to lose: medical doctors, academic researchers, taxpayers, funding organizations — and, most of all, patients. For further relevant literature, see http://meded.ucsd.edu/adpst/media_ps.html.

Ajit Varki

School of Medicine, University of California, San Diego, La Jolla, California 92039-0687, USA

Other signatories of this letter:

Edward Holmes University of California, San Diego, USA

Tadataka Yamada GlaxoSmithKline, Pennsylvania, USA

Peter Agre Duke University School of Medicine, North Carolina, USA

Sydney Brenner The Salk Institute, San Diego, California, USA

India’s concern about both security and sea research

SIR — I share marine researchers’ feelings about restrictions on carrying out research in Indian waters, as expressed in your News story “India’s ban on foreign boats hinders tsunami research” (*Nature* **439**, 380; 2006). But few countries allow foreign vessels into their ‘territorial waters’, 12 nautical miles from the coast, for research purposes. Foreign vessels are usually allowed into their exclusive economic zones, which lie between 12 and 200 nautical miles from shore.

France has made considerable efforts to

develop a joint Indo-French research programme since the 2004 tsunami. Although strongly recommended by an Indian expert panel, the programme has not been pursued. However, India has sent scientists to participate in the 2005 French marine survey, and plans to send more in July and August 2006, as part of the Sumatra-Andaman Great Earthquake research initiative (www.ipgp.jussieu.fr/~singh/SAGER).

HMS Scott, a British Royal Navy vessel, did the first marine survey off the shore of Sumatra after the tsunami. But the international science community has so far had limited access to the collected data. One should not be surprised that India is concerned about security issues. Further, India has acquired a significant amount of marine data around the Andaman-Nicobar region, both before and after the tsunami. If India develops its own marine research programme, efforts should be made to integrate these data with others recently acquired in the Indonesian waters.

Satish Singh

Laboratoire de Géosciences Marines, Institut de Physique du Globe de Paris, 4 Place Jussieu, 75252 Paris Cedex 05, France

Tools needed to navigate landscape of the genome

SIR — I enjoyed your News Feature “The web-wide world” (*Nature* **439**, 776–778; 2006), highlighting the impact of Google Earth on the scientific community. The success of this program underscores the importance of open standards for data and easy interoperation between information resources.

To some degree, the same success has been achieved in the macromolecular-structure world, where visualization tools allow users to ‘fly through’ macromolecules in 3D, and add layers of information to them. The first of these was Alywn Jones’s Frodo (T. A. Jones *J. Appl. Crystallogr.* **11**, 268–272; 1978), which inspired a generation of structural biologists. It was followed by a plethora of programs including O, Midas, Chime, Rasmol, VMD and PyMOL (see www.umass.edu/microbio/rasmol/history.htm). Many of these require only modest computing resources, and information such as mutation patterns can be easily superposed onto a structure of interest.

A good visualization program and open ‘browsing’ system can catalyse a lot of interesting science. I urge the development of comparable, free browsers for other emerging areas in the biological sciences: in particular, for visualizing the vast landscape of the genome and for navigating through complex biological networks.

Mark Gerstein

Molecular Biophysics and Biochemistry, and Computer Science, Yale University, PO Box 208114, New Haven, Connecticut 06520-8114, USA

COMMENTARY



Is it safe to put the cat
among the chickens?

© J. HABEL/CORBIS

Feline friend or potential foe?

What role do cats play in the epidemiology of H5N1 avian flu virus? We don't yet have all the answers, but it's time to consider new precautions, argue **Thijs Kuiken, Albert Osterhaus, Peter Roeder** and their colleagues.

There are increasing numbers of reports from Asia and Europe of domestic cats dying from avian influenza H5N1 virus. The available evidence, albeit incomplete, suggests that cats are more than collateral damage in avian flu's deadly global spread and may play a greater role in the epidemiology of the virus than previously thought. Here we recommend that new precautions are taken by nations and agencies fighting avian flu to minimize the risk of cats becoming infected and spreading the highly pathogenic virus to poultry, humans and other species.

The first report of domestic cats dying from H5N1 virus infections was in February 2004, when 14 out of 15 cats in a household near Bangkok, Thailand, became weak, started vomiting and coughed up blood before dying. One of the cats had eaten a chicken carcass on a farm where there was an H5N1 virus outbreak. The presence of H5N1 virus was confirmed in three of the cats following necropsies at Kasetsart University in Bangkok, Thailand¹. Subsequent experiments in The Netherlands, by some of us, have experimentally confirmed

the severity of H5N1 virus infections in cats² — with all eight cats exposed to the virus going on to develop the disease.

The jury's out

The 2004 outbreak in domestic cats showed striking similarities with an incident that occurred three months previously at a zoo in Suphanburi, Thailand, during a local outbreak of H5N1 virus infection in poultry. At this zoo, two tigers and two leopards died suddenly after feeding on fresh chicken carcasses. The probable cause of death was diagnosed as severe pneumonia due to H5N1 virus infection³. Also in 2004, there was a second outbreak of H5N1 virus infection at another zoo in Thailand, again involving consumption of virus-infected chicken. This time, a total of 147 tigers died or were killed⁴. These reports were surprising because both domestic cats and wild felids were considered to be resistant to disease from influenza A virus infection, of which H5N1 is a subtype.

Despite these unexpected events, the possible role of cats in the epidemiology of H5N1

virus infection has been largely overlooked by the human- and animal-health communities. As recently as 28 February 2006, a World Health Organization (WHO) press release⁵ stated: "There is no present evidence that domestic cats play a role in the transmission cycle of H5N1 viruses." A March press release⁶ from the World Organisation for Animal Health (OIE) stated: "The OIE stresses that as of today, all the natural cases in felines have not led to any change in the epidemiology of the disease that has fundamentally remained a bird disease, nor have they led to any recognized virus change in epidemiology or mutation leading to an increased virulence of the virus for felines or other mammals."

There are now several observations indicating a greater role for domestic cats than these cautious statements suggest. First, it has become evident that fatal infections among cats are common in countries such as Indonesia, Thailand and Iraq, where the virus seems to be endemic in poultry. For example, veterinarians conducting an epidemiology programme with the participation of the Food

and Agricultural Organization (FAO), recently reported a high incidence of sudden death among cats during fatal poultry outbreaks in several villages in Indonesia. The disease in cats is well-recognized by poultry keepers and is considered to be linked to the poultry disease. It is sufficiently well-known to have been given an onomatopoeic name in the local Javanese dialect — ‘aargh-plop’⁷.

Second, veterinarians in Iraq combating H5N1 virus outbreaks in poultry in collaboration with the FAO also report widespread and high mortality in cats, confirmed to have been caused by this virus. Third, dead or moribund cats were found to be infected with H5N1 virus soon after the virus was detected in wild birds in Germany. This suggests not only that H5N1 virus can be transmitted from wild birds to cats, but also that unusual mortality of cats in areas at risk of H5N1 virus infection may act as an early warning signal for the virus. Given the high number of infected cats in these areas, and considering their ability to excrete virus into their surroundings in sufficient quantities that transmission between cats takes place under both natural and experimental conditions (see below), cats could be more than a dead-end host for H5N1 virus.

An incomplete picture

What scientific data on H5N1 virus infection in cats are currently available? There are so far three publications^{2,8,9} reporting experimental studies, all conducted at the Erasmus Medical Centre in Rotterdam, The Netherlands. Laboratory cats infected with H5N1 virus^{2,8} by the respiratory tract (three in total), by feeding on virus-infected chicks (three), or through close contact between infected and non-infected cats (two) all excreted virus from the pharynx, nose and rectum. Excretion lasted from three days post-infection until the end of the experiment on day seven, when the cats were killed. The amount of excreted virus recorded from cats was much lower than the levels excreted by chickens.

Clinical signs were observed several days earlier in cats infected by direct respiratory and oral routes, than by cat-to-cat transmission. The signs included raised body temperature, decreased activity, conjunctivitis and laboured breathing. One of the cats died at day six. In some cats, virus excretion started before clinical signs were noted. Post-mortem examination demonstrated multiplication of the virus not only in the respiratory tract², but also in a number of other organs⁸. The presence of virus in most tissues was associated with cell death and inflammation. Viral infection in the nervous tissues of the intestinal wall was found only in cats fed on virus-infected chicks. This suggests that the virus reached these tissues directly from the guts — a novel route of entry for influenza virus in mammals.

Another recent paper of ours⁹ shows that H5N1 virus attaches abundantly to the cells lining the lungs of domestic cats, but not to those in upper parts of the respiratory tract. This pattern of H5N1 virus attachment closely mirrors that in the human respiratory tract, identifying the cat as a suitable model for viral pneumonia caused by H5N1 in humans. Some of the above findings, such as infection by feeding on virus-infected chicken carcasses, transmission of H5N1 virus between cats and infection of extra-respiratory tissues, were also observed in the H5N1 virus outbreaks in tigers and leopards and in a naturally infected domestic cat^{3,4,10}.

Although the above data shed some light on the epidemiology, major gaps in our knowledge still remain. Most important, we do not know the minimal infectious dose for cats by either the oral or the respiratory route, how long cats excrete virus from the digestive and respiratory tracts, whether cats can excrete virus without developing clinical signs, and whether virus transmission from cats to poultry, humans and other species is possible. In the absence of these data, it is difficult to assess the overall risk posed by infected cats.

Apart from the role that cats may play in H5N1 virus transmission to other species, they also may be involved in helping the virus to adapt to efficient human-to-human transmission. H5N1 virus is basically an avian pathogen, and although humans can be infected, evidence for further spread of the virus among humans is rare. In the few suspected cases¹¹, it seems that close physical contact, for example between mother and child without the use of precautions such as gloves, nose-mouth masks, or antiviral treatment, was responsible. We have previously suggested that cats may provide the virus with an opportunity to adapt to efficient transmission within and among mammalian species, including humans, thereby increasing the risk of a human influenza pandemic². Although our preliminary data from experimentally infected cats do not reveal important mutations in excreted viruses, at least within the haemagglutinin gene (the ‘H’ in H5N1), such mutations cannot be ruled out.

Time for action

Collectively, the data so far show that domestic cats can become infected by contact with domestic or wild birds and possibly their droppings, develop severe to fatal disease, excrete the virus from the respiratory and digestive tracts, and transmit the infection to other cats. Consequently, these data indicate a possible role for cats in the epidemiology of H5N1 virus in poultry, humans and other species.

Despite the uncertainties, we believe that the

potential role of cats should be considered in official guidelines for controlling the spread of H5N1 virus infection. Most international guidelines currently lack such considerations. In areas where H5N1 virus has been detected in either poultry or wild birds, we recommend taking steps to prevent contact between cats and infected birds or their droppings, and to quarantine and test cats suspected of such contacts, or cats showing clinical signs suggestive of H5N1 influenza. In most urban areas and in temperate climate zones, where most felids are domestic cats, prevention of contact could be achieved by keeping cats indoors. In other parts of the world, such measures may be more difficult, if not impossible, to implement.

Finally, we now know that H5N1 virus has the ability to infect an unprecedented range of hosts, including carnivores. In addition to felids, we can expect other domestic and wild carnivores, such as dogs, foxes, mustelids and seals, to be vulnerable to infection with H5N1 virus and to contribute to its epidemiology. Accordingly, we recommend increasing surveillance in areas where H5N1 virus is endemic, to include testing for H5N1 virus infections in felids and other carnivores showing unusual morbidity or mortality.

Only last month, H5N1 virus infection was detected in Germany in yet another mammalian species: a stone marten, belonging to the family Mustelidae. A nocturnal predator, the marten is presumed to have acquired its infections after feeding on a diseased bird. A

WHO press release responding to this report states: “To date, only domestic poultry are known to have played a role in the transmission cycle of the virus from animals to humans.” But given the potential contribution of these carnivore hosts to both virus

transmission and its adaptation to mammals, we believe the time for increased surveillance and precaution is here. ■

Thijs Kuiken, Ron Fouchier, Guus Rimmelzwaan and Albert Osterhaus are in the Department of Virology, Erasmus Medical Centre, Rotterdam, The Netherlands.

Peter Roeder is at the Animal Production and Health Division of the Food and Agriculture Organization of the United Nations, Viale delle Terme di Caracalia, Rome, Italy.

“The potential role of cats should be considered in official guidelines for controlling the spread of H5N1 virus.”

1. www.promedmail.org; archive 20040221.0560 (2004).
2. Kuiken, T. et al. *Science* **306**, 241 (2004).
3. Keawcharoen, J. et al. *Emerg. Infect. Dis.* **10**, 2189–2191 (2004).
4. Thanawongnuwech, R. et al. *Emerg. Infect. Dis.* **11**, 699–701 (2005).
5. World Health Organization www.who.int/csr/don/2006_02_28a/en/print.html (2006).
6. World Organisation for Animal Health www.oie.int/eng/press/en_060301.htm (2006).
7. Butler, D. *Nature* **440**, 135 (2006).
8. Rimmelzwaan, G. et al. *Am. J. Pathol.* **168**, 176–183 (2006).
9. van Riel, D. et al. *Science* published online doi:10.1126/science.1125548 (2006).
10. Songserm, T. et al. *Emerg. Infect. Dis.* **12**, 681–683 (2006).
11. Beigel, J. H. et al. *N. Engl. J. Med.* **353**, 1374–1385 (2005).

BOOKS & ARTS

The making of geology

In the late eighteenth century, ideas about the age of rocks and fossils gave rise to a new science.

Bursting the Limits of Time: The Reconstruction of Geohistory in the Age of Revolution

by Martin J. S. Rudwick

University of Chicago Press: 2005. 840 pp. \$45

Stephen Moorbath

Bursting the Limits of Time is a massive work and is quite simply a masterpiece of science history. It charts in the most scholarly detail the beginnings of modern geology from the mid-1780s to the mid-1820s. Much of the activity during the earlier years took place in France and adjacent countries, and the book abounds in descriptions of the ideas and interpretations of savants, including Georges Cuvier, Jean-André de Luc, Horace-Bénédict de Saussure and Abraham Werner. These scientists were beginning to show how describing and classifying geological phenomena could lead to detailed geohistorical reconstruction.

Things were slower to get going in Britain, where the most familiar names are James Hutton, William Smith and William Buckland. The book doesn't go as far as Charles Lyell, Adam Sedgwick or Charles Darwin (who was a highly productive geologist before being side-tracked by evolutionary biology), but Martin Rudwick promises to deal with the next stage of geology in a follow-up volume.

In the late eighteenth century, geology — a term that had been creeping in since the 1770s — was a complex mixture of geothory, geohistory and geognosy (an early version of structural geology) in which much of the reasoning seems remote to modern eyes. One could even say that the sheer scale of all this theorizing was inversely proportional to the amount of data available. There were occasional flashes of insight, such as Georges Buffon's view of Earth as a cooling globe with a hot centre, or Hutton's interpretation of granite as once-molten fluid. But most of the intense geothoretical arguments of those days have long since been lost or won. Now we have our own divergent geothories, which weren't on the horizon 200 years ago. But right from the start, people realized that geology, like any other science, could advance only through informed and rational disagreement.

A wind of fresh foreign air blew through the drawing rooms of the French *savanterie* with the pioneering work of Smith, the so-called father of English geology. He correlated sedi-



Rock star: Georges Cuvier realized from fossils that the Earth was much older than humans.

mentary rocks by means of their characteristic fossil assemblages, recognized the relative time sequence of sedimentary strata, made the first meaningful geological maps of parts of England and Wales, and created the important subject of stratigraphical geology, also known as stratigraphical palaeontology.

Rudwick gives Smith his full credit but points out that Alexandre Brongniart and Cuvier were doing almost identical mapping in the Paris basin at the same time and, moreover, treated fossils not only as means to a stratigraphical end, but as palaeobiological ends in themselves. It was Cuvier, with his knowledge of comparative anatomy, who recognized that most fossil bones were different from modern ones and postulated that, in the absence of fossil human bones, a long pre-human history was followed by large-scale extinction. This had been hinted at previously, and the ensuing debate centred around whether or not the supposed catastrophic event could be correlated with the biblical flood. This was a particularly important point for Buckland, who was keen to combine the broadening geological horizon with his religious perspectives.

Many of the scientists discussed their geological ideas in terms of past catastrophes and extinctions, whereas others preferred to think of vast, uneventful stretches of time, as expressed in the implicitly eternalistic words of Hutton: "The result of our present enquiry is that we find no vestige of a beginning, no

prospect of an end." In 1809, Lamarck published his gradualist 'transformation' version of evolution, quite different from that proposed by Darwin 40 years later.

As promised by its title, Rudwick's book deals expertly with the topic of geological time, which developed when no absolute measure of time was available. From studies of rocks and fossils, most scientists felt that at least the pre-human timescale was vastly longer than the 6,000 years prescribed by the Bible. In those days, theology was still regarded as related to science and was thus subject to corroboration or conflict. Scientists and theologians were in close debate over these issues, without major argument. For some today, science and religion continue to coexist peacefully, but in the early days one can see little precedent for the headline, anti-scientific attitudes practised today by fundamental biblical creationists.

Rudwick's text is beautifully written and grips the attention throughout. Nonetheless, it tends to be rather repetitive, because the author describes every nuance of each idea, however implausible, as well as the personal and scientific relationships between the protagonists. Some 200 years from now, our own scientific findings may well take just as much explaining. But a much shorter text could have covered the same subject matter, even over a broader time range, and thus reached a wider audience. All ten chapters are beautifully illustrated, with 179 contemporary black-and-white prints interwoven with the text. There are copious footnotes on every second page, as well as complete sources and references.

The book should be obligatory for every geology and history-of-science library, and is a highly recommended companion for every civilized geologist who can carry an extra 2.4 kg in his rucksack.

Cuvier ultimately emerges as the pivotal figure in the history traced in this book, but numerous others of scarcely less importance produced the foundations of the new geology. Rudwick has amply fulfilled his stated aim of describing the injection of history into a science that had been primarily descriptive or causal. Indeed, thanks to Rudwick and his kind, we may rest assured that the future of the history of science is in safe hands. ■

Stephen Moorbath is in the Department of Earth Sciences, University of Oxford, Oxford OX1 3PR, UK.

Men of astronomy

Revolutionaries of the Cosmos: The Astro-Physicists

by Ian Glass

Oxford University Press: 2006. 336 pp.
\$74.50, £35

J. D. Fernie

As a youth, Ian Glass was inspired by Eric Temple Bell's book *Men of Mathematics* (Simon & Schuster, 1937), which profiled more than 30 mathematicians. In *Revolutionaries of the Cosmos*, Glass has attempted to do much the same for astronomy. He has restricted himself to just eight subjects, and has clearly been careful in his biographical research, which he documents in detail after each chapter. There are about 30 or 40 pages per individual, which is an admirable length, as one often wants more than a brief dictionary entry but less than a full-scale biography.

The selection of individuals — Galileo, Isaac Newton, William Herschel, William Huggins, George Ellery Hale, Arthur Eddington, Harlow Shapley and Edwin Hubble — may raise a few eyebrows. Glass chose them because each “made at least one important discovery by applying the methods of physical science to astronomy” — hence the reference in the subtitle to ‘astro-physicists’, with the hyphen emphasizing the historical development. Even so, the methods of physical science changed considerably over the centuries and it seems somewhat unbalanced that a book called *Revolutionaries of the Cosmos* contains only passing mentions of Tycho Brahe and Johannes Kepler, yet Huggins receives the full treatment in recognition of his spectroscopic achievements. And it seems that Herschel, despite his telescopic discoveries, believed for most of his life that the Sun and Moon were inhabited by living creatures. That all these people, with the exception of Galileo, came from the English-speaking world means that the book is not a balanced history of astronomy, but that obviously was not the author's intent. Researchers from non-English-speaking nations often enter the story in passing, however, providing some balance.

Glass writes clearly, interestingly and evenhandedly throughout. His opening chapters cover material that will be familiar to most readers, but he brings out details that many might be unaware of, such as the fact that Galileo received considerable support from the more progressive church officials in his confrontation with the Catholic Church. And that while Galileo held the mathematics chair at Padua University, an official mooted an increase in salary for him because he had acquired a mistress.

But it is the later chapters, mainly covering the twentieth century, that will appeal to most readers, especially those who have at least a casual knowledge of astronomical history in

The Hale telescope, dedicated in 1948, commemorates one of the founders of modern astrophysics.



BETTMANN/CORBIS

this period. The story of Hale is particularly absorbing: he founded two of the world's most important observatories, as well as the California Institute of Technology, and set up some of the world's largest telescopes, yet he battled mental-health problems for much of his life.

The chapter on Eddington is similarly fascinating. He was attacked by other British scientists around the time of the First World War because, as a Quaker, he argued that most German scientists were not the monsters they were made out to be. His pacifist views nearly led to his imprisonment by the British government. Glass also discusses Eddington's interactions with his rival, James Jeans. Their violent arguments over each other's papers presented at meetings of the Royal Astronomical Society were so intense that some people, such as G. H. Hardy, joined the society just to have a ringside seat! Yet it was a common sight to see Eddington and Jeans enjoying afternoon tea together.

Glass's writing includes quiet touches of humour here and there. For example, he tells us that early prospective authors writing for *The Astrophysical Journal* had to face its three editors: Hale, Gale and Frost. And that the budget for building the world's finest spectroheliograph in 1899 was so small that one of the lenses had to be bought from a pawnbroker.

Although Glass doesn't comment on it, his book does implicitly raise the question of whether or not the future will be very different. One has only to look at *The Astrophysical Journal* now to realize that the day of the lone researcher is over: large teams are the order of the day, and it's doubtful whether individuals will ever be as influential as they used to be. Meanwhile, Glass has written an absorbing book, which I strongly recommend. ■

J. D. Fernie is emeritus professor in the Department of Astronomy and Astrophysics, University of Toronto, 60 St George Street Toronto, Ontario M5S 3H8, Canada.

The puzzle of cooperation

Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life

edited by Herbert Gintis, Samuel Bowles, Robert Boyd & Ernst Fehr

MIT Press: 2005. 404 pp. \$50. £32.95

Andrew M. Colman

Robert May began his last presidential address to the Royal Society on 30 November 2005 by saying: “The most important unanswered question in evolutionary biology, and more generally in the social sciences, is how cooperative behaviour evolved and can be maintained in human or other animal groups and

societies”. For example, birds often emit alarm calls when they spot predators, but how could such behaviour have evolved? A mutant bird that never gave alarm calls would save energy and avoid the additional risk to itself while enjoying the benefits of its conspecifics' alarm calls. Its ‘selfish gene’ should therefore spread to fixation in the population.

For the same reason, cooperation is difficult to maintain when individuals are tempted to defect. A recent human example in Britain is the decline in voluntary take-up of the combined measles–mumps–rubella (MMR) vaccination by parents who wished to avoid an alleged health risk to their own children while

implicitly relying on enough other children being vaccinated to maintain 'herd immunity'. This has the strategic structure of a social dilemma, because if all parents followed this reasoning, everyone could end up worse off than if they all behaved cooperatively.

A similar social dilemma is devastating UK fish stocks: overfishing destroyed British herring fisheries long ago and is now causing terminal decline in other fish stocks in the English Channel, the North Sea and the Baltic. Anyone who makes a living by fishing is motivated to catch as many fish as possible, because restraint is pointless if enough others are exercising restraint, and is futile if they are not. But then fish are driven to extinction and everyone is worse off than if they had all restrained themselves cooperatively.

Bill Hamilton's theory of 'inclusive fitness', or kin selection, explains the evolution of cooperation among genetically related individuals. It can explain the extreme self-sacrificing cooperation of female social Hymenoptera, who have 75% of their genes in common, but cannot explain cooperation among non-relatives. Robert Trivers's theory of reciprocal altruism shows how cooperation between non-relatives can evolve if its cost is small and is outweighed by favours returned in the future. These two theories go some way towards explaining how cooperation evolved, but neither can explain human cooperation in unrepeated interactions between strangers.

To fill this gap, Ernst Fehr and Simon Gächter introduced in 2000 a version of the theory of strong reciprocity incorporating the 'altruistic punishment' of non-cooperators. *Moral Sentiments and Material Interests* is devoted to their theory's biological, anthropological, economic and social ramifications and related ideas. An introductory chapter is followed by three on the behavioural ecology of cooperation, four on modelling and testing strong reciprocity, and five on reciprocity and social policy, all by researchers in the vanguard of their fields.

According to the theory, cooperation is necessary for the provision of public goods, and the punishment of non-cooperators, or free-riders, is itself a public good — a service provided for the benefit of the whole community. Such punishment is altruistic because it is costly to those who administer it, as it takes time and energy and invites retaliation. Fehr and Gächter have provided persuasive experimental evidence, reviewed in this book, that cooperation flourishes when punishment is possible and breaks down when it is not.

A problem not addressed in the book is that, because altruistic punishment is costly, natural selection should tend to eliminate it. Failure to punish defectors must presumably be treated as second-order defection, itself subject to sanctions from other group members. But what about sanctions against third-order defectors who neglect to punish second-order defectors, and so on? This is an infinite regress



Tragic commons: North Sea fishing boats sit idle, thanks to rational economic decision-making.

that becomes less credible with the addition of each successive layer. Strong reciprocity is an important and illuminating discovery, but we seem to have replaced the problem of explaining cooperation with that of explaining altruistic punishment.

The book provides a superb interdisciplinary synthesis of cooperation as explained by strong reciprocity and associated phenomena. Other explanations of cooperation in unrepeated interactions between strangers hardly get a look in, however. The most important is Richard Alexander's theory of 'indirect reciprocity',

according to which people use observations of direct reciprocity between others when deciding how to act towards them in the future. People benefit by cooperating, even in one-off encounters with strangers, because cooperation enhances one's reputation for cooperativeness and elicits reciprocal cooperation from others. This is a powerful theory, supported by evidence from computational and experimental studies, but *Moral Sentiments and Material Interests* mentions it only in passing.

Andrew M. Colman is in the School of Psychology, University of Leicester, Leicester LE1 7RH, UK.

Back to the drawing board?

Knowing: The Nature of Physical Law

by Michael Munowitz

Oxford University Press: 2006. 432 pp

£19.99, \$35

David Lindley

Among the many services that Michael Faraday rendered to science was his innovation of illustrating theoretical ideas by means of pictures and diagrams. For Faraday, who knew next to no mathematics but had a powerful visual imagination, the image was the concept. But as his successors dressed his ideas in the finery of nineteenth-century mathematics, we have come to think of mathematics as the essence of theory, with pictures being merely a helpful adjunct. Attempting to push the pendulum back a little, Michael Munowitz has written an account of modern physics that Faraday might appreciate, using no mathematics (well, all right, there is some arithmetic) but plenty of illustrations. His mixed success shows, rather sadly perhaps, how far theory has moved beyond our ability to capture it in an easily grasped visual form.

Munowitz plunges in with an opening chapter devoted to the four elementary forces of nature. Gravity holds the Solar System together,

and electromagnetic forces do much the same thing for atoms. Going to a still smaller scale, we see how there must be a strong nuclear force to prevent positively charged protons from bursting out of the nucleus. This is nicely done. In each case, we perceive the force through its ability to control a physical system.

But then comes the pesky weak nuclear force, or rather interaction. The change of word betrays the problem. If there is some visualizable structure that the weak interaction holds together, I can't think of it, and neither can Munowitz. Instead, he talks about radioactive decays that the weak interaction engenders — and the connection to something that might be called a force is quietly dropped. I don't mean this as a criticism so much as a recognition of how difficult a task Munowitz has set himself.

Where Munowitz succeeds admirably, however, is in his recurrent emphasis on the way that principles of invariance and symmetry dictate the shape of natural law. Insistence that physics cannot depend on absolute velocity — there is no such thing — leads to the conservation law for momentum. Likewise, the impossibility of absolute time takes us (not so obviously) to the conservation of energy. In neither case can Munowitz make

the connection precise, but he manages to make the reader feel that the conservation laws are not *ex cathedra* pronouncements but derive from eminently sensible assertions about the way the Universe is constructed.

A later chapter builds on these arguments to show how local gauge invariance of the electromagnetic field, for example, demands a force to make sure that particles behave as they must towards each other. And the tautological principle that an observer cannot tell the difference between identical quantum objects leads into a nice account of Pauli's principle

and exchange forces. In these sections Munowitz splendidly demonstrates how a handful of bedrock principles underlie many apparently different areas of physics. The avoidance of mathematics is a real advantage here, forcing attention on to broad concepts rather than specific rules.

From time to time I had the nagging feeling that this book may be illuminating in the way that Richard Feynman's celebrated lectures are — by striking flashes of intellectual enlightenment in readers who have already learned physics the old-fashioned way. A young Fara-

day of our times, looking to this book for an education, would in the end be obliged to take many of the conclusions on trust. And in a few chapters, Munowitz fumbles the conceptual continuity, delivering catalogues of things that the dutiful reader ought to know. But for any high-school student or undergraduate who is losing sight of the forest on account of all the trees, *Knowing* at times offers a persuasively harmonious view of the physicist's way of looking at the world. ■

David Lindley is a freelance writer based in Alexandria, Virginia, USA.

High impact

From protons to galaxies, *Cosmic Collisions* shows us what happens when things go bump.

Michael Hopkin

Our galaxy is big. But it's due to get a lot bigger: in a few billion years, the Milky Way will slam into its nearest neighbour, the Andromeda Spiral, eventually creating a super-sized stellar clump. And visitors to the Hayden Planetarium at New York's American Museum of Natural History can now see a preview of the monumental event as it is likely to unfold.

The scene is the denouement of *Cosmic Collisions*, a \$3-million film that reveals the power of impacts great and small to shape

our world and the Universe. The project also showcases the prodigious supercomputing power now available to astrophysicists — the largely computer-generated movie visualizes some of theorists' most advanced simulations of cosmic processes.

The film focuses on impacts across a mind-boggling range of scales, from the proton collisions that power the Sun, to the asteroid smash that may have done for the dinosaurs, to the *pas de deux* of galaxies played out over millions of years.

The team behind the film, including the

museum's astrophysics department, NASA and more than 25 researchers from around the world, based their pictures on supercomputer simulations and satellite images.

The starry sky, created using a virtual map called the Digital Universe, forms a backdrop to the film's opening, over which the gravelly tones of Robert Redford inform us that, although space might look tranquil, its myriad collisions make it a violent place.

This becomes clear during a scene in which the serenely drifting Earth is suddenly slammed by a planet-sized body that seems to appear from nowhere. This event really happened, 4.5 billion years ago, and in the space of a month it gave us our Moon. The film draws on simulations of lumps of debris being drawn together by gravity to show us how this remarkably rapid process occurred.

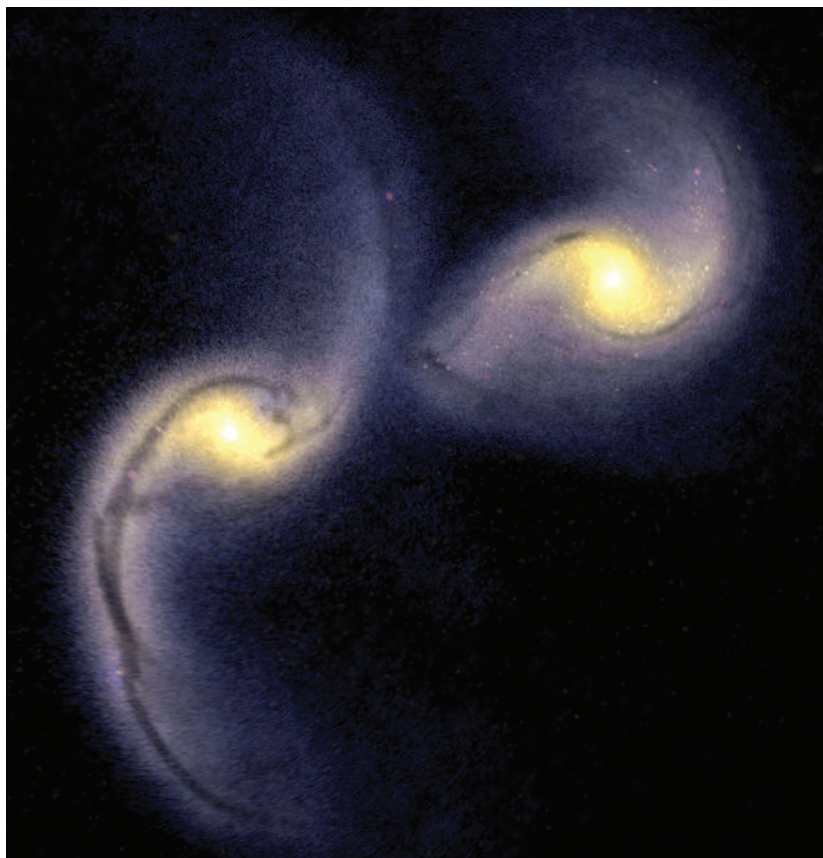
The film culminates with the climactic sequence in which two galaxies, each containing hundreds of billions of stars, meet. The sequence, which lasts less than a minute but in which every second represents 40 million years, is the product of more than 40,000 hours of computing time.

The galactic clash is based on simulating the behaviour of individual stars within each cluster, and comparing the results with data on real galactic collisions. Transforming that information into a visual sequence was the work of yet more supercomputers.

As for the event itself, the collision won't cause much disruption to any civilizations still around. As Redford reassures us, "Stars and planets in these galaxies won't actually collide. They're much too far apart. Scientists think they'll simply slide past one another." Nevertheless, you'll have to go a lot further than New York to get such a grandstand view of the real thing.

Cosmic Collisions is showing at the Hayden Planetarium, American Museum of Natural History, New York.

Michael Hopkin writes for news@nature.com.



Forthcoming attraction: a simulation of the Milky Way's merger with Andromeda (left).

NEWS & VIEWS

PALAEOONTOLOGY

A firm step from water to land

Per Erik Ahlberg and Jennifer A. Clack

A project designed to discover fossils that illuminate the transition between fishes and land vertebrates has delivered the goods. At a stroke, our picture of that transition is greatly improved.

The concept of 'missing links' has a powerful grasp on the imagination: the rare transitional fossils that apparently capture the origins of major groups of organisms are uniquely evocative. But the concept has become freighted with unfounded notions of evolutionary 'progress' and with a mistaken emphasis on the single intermediate fossil as the key to understanding evolutionary transitions. Much of the importance of transitional fossils actually lies in how they resemble and differ from their nearest neighbours in the phylogenetic tree, and in the picture of change that emerges from this pattern.

We raise these points because on pages 757 and 764 of this issue^{1,2} are reports of just such an intermediate: *Tiktaalik roseae*, a link between fishes and land vertebrates that might in time become as much of an evolutionary icon as the proto-bird *Archaeopteryx*. Several specimens have been found in Late Devonian river sediments on Ellesmere Island in Nunavut, Arctic Canada. They show a flattened, superficially crocodile-like animal, with a skull some 20 centimetres in length. The body is covered in rhombic bony scales, and the pectoral fins are almost-but-not-quite forelimbs; these contain robust internal skeletons, but are fringed with fin rays rather than digits. *Tiktaalik* goes a long way — but not quite the whole way — towards filling a major gap in the picture of the vertebrate transition from water to land.

It has long been clear that limbed vertebrates (tetrapods) evolved from osteolepiform lobe-finned fishes³, but until recently the morphological gap between the two groups remained frustratingly wide. The gap was bounded at the top by primitive Devonian tetrapods such as *Ichthyostega* and *Acanthostega* from Greenland, and at the bottom by *Panderichthys*, a tetrapod-like predatory fish from the latest Middle Devonian of Latvia (Fig. 1). *Ichthyostega*⁴ and *Acanthostega*⁵ retain true fish tails with fin rays but are nevertheless unambiguous tetrapods with limbs that bear digits⁶. *Panderichthys*⁷ is vaguely crocodile-shaped and, unlike the rather conventional osteolepiform fishes farther down the tree, looks like a

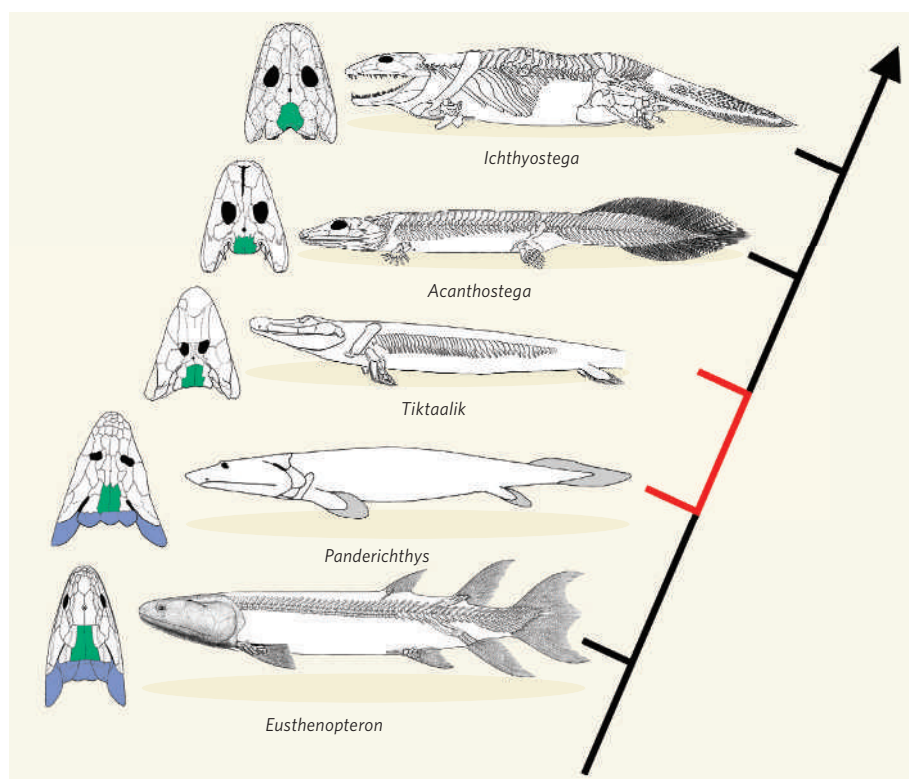


Figure 1 | *Tiktaalik* in context. The lineage leading to modern tetrapods includes several fossil animals that form a morphological bridge between fishes and tetrapods. Five of the most completely known are the osteolepiform *Eusthenopteron*¹⁶; the transitional forms *Panderichthys*¹⁷ and *Tiktaalik*¹; and the primitive tetrapods *Acanthostega* and *Ichthyostega*. The vertebral column of *Panderichthys* is poorly known and not shown. The skull roofs (left) show the loss of the gill cover (blue), reduction in size of the postparietal bones (green) and gradual reshaping of the skull. The transitional zone (red) bounded by *Panderichthys* and *Tiktaalik* can now be characterized in detail. These drawings are not to scale, but all animals are between 75 cm and 1.5 m in length. They are all Middle–Late Devonian in age, ranging from 385 million years (*Panderichthys*) to 365 million years (*Acanthostega*, *Ichthyostega*). The Devonian–Carboniferous boundary is dated to 359 million years ago¹⁸.

fish–tetrapod transitional form. The shape of the pectoral fin skeleton and shoulder girdle are intermediate between those of osteolepiforms and tetrapods, suggesting that *Panderichthys* was beginning to 'walk', but perhaps in shallow water rather than on land⁸.

Panderichthys lived about 385 million years ago at the end of the Middle Devonian; *Ichthyostega* and *Acanthostega* lived about 365 million years ago during the Late Devonian. However, the earliest fragmentary tetrapods from

Scotland^{9,10} and Latvia⁹ date back to perhaps 376 million years ago, so the morphological gap between fish and tetrapod corresponds to a time gap of under 10 million years.

Into this gap drops *Tiktaalik*. The fossils are earliest Late Devonian in age, making them at most 2 million or 3 million years younger than *Panderichthys*. With its crocodile-shaped skull, and paired fins with fin rays but strong internal limb skeletons, *Tiktaalik* also resembles *Panderichthys* quite closely. The closest match,

EVOLUTION

It pays to laze

Hidden beneath small mounds in the Kalahari Desert in southern Africa, Damaraland mole-rats (*Cryptomys damarensis*, pictured) have developed a remarkable caste system. In the life cycle of these animals, which is spent entirely underground, a single 'queen' female mates with one or two unrelated males. The rest of the colony members generally invest their efforts in caring for successive litters of young, hunting for food and maintaining the colony's intricate network of tunnels.

These worker mole-rats are divided into two types: 'frequent' and 'infrequent' workers, the latter being evidently lazy types that may comprise as much as 40% of the community but do less than 5% of

the work. Elsewhere in this issue, M. Scantlebury *et al.* describe how they have followed up circumstantial evidence for the reasons behind this division of labour, and show that in certain situations the layabouts spring into action (*Nature* **440**, 795–797; 2006).

Mole-rat workers are thought to postpone their own reproduction (sometimes indefinitely) because of the difficulties of setting up a new colony in the rock-hard soil. Extensive burrowing, and so the chance of meeting a mate from another colony, is restricted to brief periods, maybe once or twice a year, when heavy rains soften the soil.

Is this when infrequent workers pay their dues? To find out, Scantlebury and colleagues examined individuals they trapped at burrow entrances. By measuring the body fat, daily energy

expenditure and resting metabolic rate of several individuals during a dry period and after rainfall, the authors show that the infrequent workers are fatter and expend far less energy than the other workers when it is dry. Following rainfall, however, they display bursts of effort not shown by the other colony members. Scantlebury *et al.* propose that, by conserving their energy during dry periods and then digging furiously after it has rained, the fat workers have a good chance of dispersing far enough to find a mate.

As the authors point out, funnelling extra resources into a dispersive caste may well be a sensible strategy for the colony as a whole. These apparent layabouts may spend most of their time reaping the benefits of colony life, such as food and protection, without pulling their weight. But they seem



T. JACKSON/CERU/UNIV. PRETORIA

to give good returns when it comes to exploiting environmental conditions to ensure long-term survival of the colony's gene pool.

Lucy Odling-Smee

however, is not to *Panderichthys* but to another animal, *Elpistostege*, from the early Late Devonian of Canada. *Elpistostege* is known only from two partial skulls and a length of backbone, but it has long been recognized as a fish–tetrapod intermediate^{11,12}, probably closer to tetrapods than is *Panderichthys*. This impression is now confirmed: the authors^{1,2} demonstrate convincingly that *Elpistostege* and *Tiktaalik* fall between *Panderichthys* and the earliest tetrapods on the phylogenetic tree.

So, if *Tiktaalik* is in effect a better-preserved version of *Elpistostege*, why is it important? First, it demonstrates the predictive capacity of palaeontology. The Nunavut field project had the express aim of finding an intermediate between *Panderichthys* and tetrapods, by searching in sediments from the most probable environment (rivers) and time (early Late Devonian). Second, *Tiktaalik* adds enormously to our understanding of the fish–tetrapod transition because of its position on the tree and the combination of characters it displays.

In some respects, *Tiktaalik* and *Panderichthys* are straightforward fishes: they have small pelvic fins¹³, retain fin rays in their paired appendages and have well-developed gill arches, suggesting that both animals remained mostly aquatic. In other regards, *Tiktaalik* is more tetrapod-like than *Panderichthys*. The bony gill cover has disappeared, and the skull has a longer snout (Fig. 1). These changes probably relate to breathing and feeding, which are linked in fishes because the movements used for gill ventilation can also be used to suck food into the mouth. A longer snout suggests a shift from sucking towards snapping up prey, whereas the loss of the gill cover bones (which turned the gill cover into a soft flap) probably

correlates with reduced water flow through the gill chamber. The ribs also seem to be larger in *Tiktaalik*, which may mean it was better able to support its body out of water¹. The only real peculiarity of *Tiktaalik* is its poorly ossified vertebral column that seems to contain an unusually large number of vertebrae.

These character distributions paint an intriguing picture. *Tiktaalik* is clearly a transitional form, more tetrapod-like than *Panderichthys* in its breathing and feeding apparatus, but with similar locomotory adaptations. Crucially, because *Tiktaalik* occupies a position closer to tetrapods on the tree than does *Panderichthys*, their shared characters can be inferred to be attributes of the segment of the tree between the branches that carry the two animals (Fig. 1, red). *Panderichthys* showed us a morphology that could be interpreted as directly intermediate between osteolepiform and tetrapod. But only the similar yet 'upgraded' morphology in *Tiktaalik* demonstrates that this interpretation is correct: this really is what our ancestors looked like when they began to leave the water.

Two aspects of *Tiktaalik*'s anatomy relate to the origin of new structures in tetrapods: the ears and limbs. The tetrapod middle ear has arisen as a modification of the fish spiracle (a small gill slit) and hyomandibula (a bone supporting the gill cover). *Panderichthys* possesses a widened spiracle, interpreted as the intake for air or water, and a shortened hyomandibula¹⁴. *Tiktaalik* shows an almost identical condition, but with an even wider spiracle, indicating that this morphology too is genuinely transitional.

The pectoral fin skeleton of *Tiktaalik* is notable not only because of its transitional nature, but also because its excellent preservation has

allowed the individual bones to be freed of the rock and manipulated to estimate ranges of movement². It turns out that the distal part of the skeleton is adapted for flexing gently upwards — just as it would if the fin were being used to prop the animal up. Although these small distal bones bear some resemblance to tetrapod digits in terms of their function and range of movement, they are still very much components of a fin. There remains a large morphological gap between them and digits as seen in, for example, *Acanthostega*: if the digits evolved from these distal bones, the process must have involved considerable developmental repatterning. The implication is that function changed in advance of morphology.

The body form represented by *Tiktaalik* and *Panderichthys* was evidently an actual step on the way from water to land. Just over 380 million years ago, it seems, our remote ancestors were large, flattish, predatory fishes, with crocodile-like heads and strong limb-like pectoral fins that enabled them to haul themselves out of the water. Further information will emerge from the full description of the fossils, and from detailed comparisons with Devonian tetrapods such as the very primitive *Ventastega*¹⁵.

Of course, there are still major gaps in the fossil record. In particular we have almost no information about the step between *Tiktaalik* and the earliest tetrapods, when the anatomy underwent the most drastic changes, or about what happened in the following Early Carboniferous period, after the end of the Devonian, when tetrapods became fully terrestrial. But there are still large areas of unexplored Late Devonian and Early Carboniferous deposits in the world — the discovery of *Tiktaalik* gives hope of equally ground-breaking finds to come. ■

Per Erik Ahlberg is in the Department of Physiology and Developmental Biology, Uppsala University, Norbyvägen 18A, 752 36 Uppsala, Sweden.

e-mail: per.ahlberg@ebc.uu.se

Jennifer A. Clack is in the Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK.

e-mail: jac18@hermes.cam.ac.uk

1. Daeschler, E. B., Shubin, N. H. & Jenkins, F. A. Jr *Nature* **440**, 757–763 (2006).
2. Shubin, N. H., Daeschler, E. B. & Jenkins, F. A. Jr *Nature* **440**, 764–771 (2006).
3. Cope, E. D. *Proc. Am. Phil. Soc.* **30**, 278–281 (1892).
4. Ahlberg, P. E. *et al.* *Nature* **437**, 137–140 (2005).
5. Coates, M. I. *Trans. R. Soc. Edinb. Earth Sci.* **87**, 363–421 (1996).
6. Clack, J. A. *Gaining Ground: The Origin and Early Evolution of*

- Tetrapods* (Indiana Univ. Press, Bloomington, 2002).
7. Vorobyeva, E. I. & Schultze, H.-P. in *Origins of the Higher Groups of Tetrapods* (eds Schultze, H.-P. & Trueb, L.) 68–109 (Comstock, Ithaca, 1991).
 8. Vorobyeva, E. I. & Kuznetsov, A. in *Fossil Fishes as Living Animals* (ed. Mark-Kurik, E.) 131–140 (Inst. Geol., Tallinn, Estonia, 1992).
 9. Ahlberg, P. E. *Nature* **373**, 420–425 (1995).
 10. Ahlberg, P. E. *J. Linn. Soc.* **122**, 99–141 (1998).
 11. Westoll, T. S. *Nature* **141**, 127–128 (1938).
 12. Schultze, H.-P. & Arsenault, M. *Palaeontology* **28**, 293–309 (1985).
 13. Boisvert, C. *Nature* **438**, 1145–1147 (2005).
 14. Brazeau, M. & Ahlberg, P. E. *Nature* **439**, 318–321 (2006).
 15. Ahlberg, P. E., Lukševičs, E. & Lebedev, O. A. *Phil. Trans. R. Soc. Lond. B* **343**, 303–328 (1994).
 16. Jarvik, E. *Basic Structure and Evolution of Vertebrates* (Academic, New York, 1980).
 17. Ahlberg, P. E. & Milner, A. R. *Nature* **368**, 507–514 (1994).
 18. Gradstein, F. M. *et al.* *A Geologic Time Scale 2004* (Cambridge Univ. Press, 2005).

SEMICONDUCTORS

Spray-on silicon

Lisa Rosenberg

Reports of the death of silicon electronics may well have been exaggerated. A technique that allows the deposition of silicon films from solution could harbinger the era of the inkjet-printed circuit.

On page 783 of this issue, Shimoda *et al.*¹ set forth a radical way of incorporating silicon into that most basic of electronic components, the transistor. Their technique uses a novel liquid precursor of solid silicon to allow the 'printing' of semiconductor films via familiar inkjet technology. It could thus permit unprecedented control over the size and placement of semiconducting silicon in future generations of high-performance electronic equipment.

When its structure is delicately disrupted, ultra-pure silicon is the quintessential semiconductor. As such, it controls the electrical impulses that in turn control the computers and other electronic devices that many of us take for granted. Semiconducting silicon is obtained from highly purified natural silicon — which occurs in the form of silica (silicon dioxide) in quartzite rock and sands — by adding tiny amounts of appropriate impurities (the process known as doping), or through specialized crystallization methods. Whichever way is chosen, enormous effort goes into extracting, refining, shaping and processing silicon to make it technologically useful.

Recently, the need for semiconducting transistors to help transform electricity into coloured light for displays and screens has begun to present a challenge for existing silicon manufacturing technologies. In particular, the demand for screens with ever-increasing pixel resolution that are thinner, brighter, wider and lighter — or even flexible — has stretched the solid-state patterning techniques used to produce silicon circuitry to the limit. It has also fuelled intense research into

alternative, more processable semiconducting materials, including organic molecules and polymers^{2–4}.

Electronic devices based on solid layers of silicon still provide the benchmark for semiconductor performance, however. Conventional techniques for their manufacture involve, for instance, heating ultra-pure silicon in a vacuum to create a mist of free silicon atoms that condenses onto a supporting surface, preferably an inexpensive plastic. But multiple refining and deposition steps are still just the start of the delicate and convoluted manufacturing process that leads to a finished transistor (see Fig. 4d on page 785). Once deposited, the solid film must be sliced or etched to produce circuit elements, and then attached to the rest of the electronic components. All this must be done without compromising silicon's semiconducting properties. That places severe limits on semiconductor thickness, patterning and connectivity, and therefore on advances in electronics design.

Controlling a liquid is generally easier than sculpting a solid. Sophisticated, high-resolution printing technologies already exist that could be used to introduce very thin layers of liquid semiconductor in complex patterns over a variety of surfaces. The advantage would be particularly great for the mass production both of very small devices and of displays that cover huge areas, as their transistor components must typically be no more than about a thousandth of a millimetre thick. In both such applications, it is difficult to consistently reproduce transistor size, thickness and pattern using solid-state techniques.

But how can silicon be produced in liquid form? Molten silicon is out of the question here: with its melting temperature of 1,414 °C, it would destroy the other components of the device, as well as the printing equipment. Shimoda and colleagues' solution¹ is to delve into the realm of decades-old synthetic chemistry. They focus on a binary compound of silicon and hydrogen, Si₅H₁₀ or cyclopentasilane, that is liquid at room temperature. When baked at a temperature of 300 °C or higher, this compound loses hydrogen gas, leaving a residue of pure, elemental silicon. It is thus, apparently, an ideal liquid precursor for silicon thin films.

But there is a problem. Cyclopentasilane tends to boil off during baking, making it difficult to control the amount of silicon that is actually left behind. The authors circumvent this obstacle using a technique called 'ring-opening' polymerization chemistry. They shine light of ultraviolet wavelength on to the five-membered silicon rings in the cyclopentasilane liquid, causing them to open and join end-to-end. The result is long, non-volatile chains, known as polysilanes, that have the characteristics of viscous oils or even solids. If this polymerization is halted part-way through, the polysilanes already produced dissolve in the remaining unconverted cyclopentasilane, resulting in a solution from which an elemental silicon residue can form. The amorphous network of silicon atoms, a-Si, obtained does not have the optimum three-dimensional structure for semiconducting behaviour, and so, as a final step, high-intensity ultraviolet light is applied to rearrange it into a more ordered, polycrystalline form (poly-Si).

Shimoda and colleagues are thus the first to produce relatively high-performance silicon films by processing from solution. They first prepared films by simple spin coating — essentially, spraying a thin layer of solution onto a quartz surface before baking — and found that the properties of the films were comparable to those of high-quality poly-Si produced by conventional techniques. Although this performance was lower for films deposited using inkjet-printing technology, it was still much higher than is typically achieved for solution-processed films based on alternative, organic materials⁴.

This method dispenses with some high-temperature refining of metallurgical silicon extracted from silica and replaces it with chemical synthesis and milder distillations. Admittedly, this liquid polysilane precursor is highly sensitive to contamination by oxygen both during and after its preparation. Such contamination can drastically diminish the electronic performance of the eventual film, and dictates that air and water must be rigorously excluded at all stages of the process. These precautions are, however, no different from those taken for traditional routes to silicon thin films.

But it is the potential for taking advantage of highly controlled printing techniques in the

patterning of silicon thin films that makes this work particularly exciting. Challenges do remain here: the physical properties of the liquid precursor, such as its viscosity, vapour pressure and surface tension, must be tuned to take full advantage of the printing tools available, while precisely controlling the post-baking thin-film morphology critical to semiconductor performance. Careful study of dilution effects is also needed. Finally, stricter control of the distribution of polysilane chain lengths within the precursor mixture would also probably help — a real test for polymer chemists.

It might be that, in the end, inkjet printing of 'liquid silicon' will not provide the resolution necessary to pattern a high-density integrated

circuit and therefore make a computer chip. But what it will certainly allow is the remarkably straightforward generation of simple, cheap and flexible circuits for displays, as well as a range of other applications — solar cells, X-ray detectors and multi-analyte chemical sensors included.

Lisa Rosenberg is in the Department of Chemistry, University of Victoria, PO Box 3065, Victoria, British Columbia V8W 3V6, Canada. e-mail: lisarose@uvic.ca

1. Shimoda, T. *et al. Nature* **440**, 783–786 (2006).
2. Horowitz, G. *Adv. Mater.* **10**, 365–377 (1998).
3. Dimitrakopoulos, C. D. *et al. Adv. Mater.* **14**, 99–117 (2002).
4. Mitzi, D. B. *et al. Nature* **428**, 299–303 (2004).

IMMUNOLOGY

The pick of the nibbled bits

Joel Swanson

How does the immune system avoid potentially damaging responses against the body's own molecules? The answer lies partly in the ability of dendritic cells to sample their surroundings selectively.

Animal tissues are groomed by phagocytes, migratory cells that routinely ingest moribund neighbouring cells, infecting microbes and particulate debris. In the immune system, specialized phagocytes called dendritic cells recycle the molecular fragments from ingested proteins, transporting the scraps to their plasma membranes for display on their external surface as antigens. If the fragments are derived from a microbe, then the dendritic cell adds indicators of the alien status of its meal onto its surface. The displayed antigens are recognized by circulating T lymphocytes, cells that can activate specific immune responses if they later find the same alien molecules presented elsewhere in the body¹. The voracious appetites of phagocytes suggest a dilemma concerning mixed meals: phagocytosis of two particles, one that can trigger an immune response and another that cannot, could launch a misguided response against benign molecules.

On page 808 of this issue, Blander and Medzhitov² show that phagosomes, the intracellular organelles that contain ingested particles and antigenic proteins, can work independently of one another inside the same cell to determine the fate of their enclosed antigens. This subcellular regulation suggests a mechanism for focusing the immune response and provides clues for the rational design of vaccines.

Immature dendritic cells in peripheral tissues 'sample' their environment by phagocytosis, initiating specific immune responses when they sense microbes or tissue damage. Ingested material is usually degraded completely

as phagosomes mature into phagolysosomes. The maturation, or 'differentiation', of dendritic cells can be stimulated in part by signalling through Toll-like receptors (TLRs), a family of cell-surface molecules that recognize molecular patterns common to microorganisms.

TLR4 recognizes lipopolysaccharide, a

conserved and abundant surface molecule of Gram-negative bacteria such as *Escherichia coli*. Through TLR4, the lipopolysaccharide ligand stimulates dendritic-cell maturation and the re-routing of the cell's intracellular traffic, such that fragments of proteins generated in phagolysosomes are bound by 'MHC class II' proteins to make antigen-presentation complexes that move to the cell surface³. T lymphocytes capable of recognizing specific peptide antigens in MHC class II complexes continually scan surfaces of mature dendritic cells, and recognition of their cognate ligand typically leads to activation of an immune response specific to that antigen (Fig. 1).

Blander and Medzhitov² examined whether TLR4-mediated differentiation affects the fate of all phagosomes in the cell or only those containing the TLR4 ligand. They allowed dendritic cells from mice to ingest bacteria or dying (apoptotic) cells containing protein antigens, such that antigen processing could occur in the presence or absence of phagosome-restricted TLR4 signalling. They then compared the antigen-presentation responses when antigen and lipopolysaccharide were eaten together either on the same particle or in distinct particles inside the same cell.

Protein antigens delivered via *E. coli*, which would be expected to activate TLR4, were fully processed into peptide-loaded MHC class II complexes and were presented to antigen-specific T lymphocytes at the dendritic-cell surfaces. However, proteins delivered by phagocytosis of apoptotic cells, which lack TLR4 ligands, were not presented to T lymphocytes, even when those same dendritic

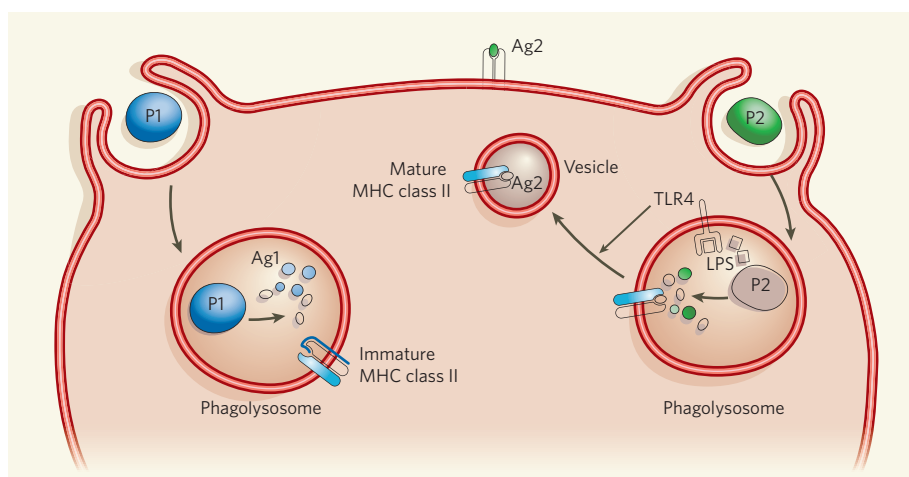


Figure 1 | Independence for phagosomes. Dendritic cells ingest detritus from their environment into organelles called phagosomes, where particle-associated proteins (P1 and P2) are partially degraded into antigenic fragments (Ag1 and Ag2). Phagosomes mature into phagolysosomes. If the antigens come from proteins that are alien to the body (P2), the antigens (Ag2) associate with MHC class II antigen-presenting complexes, and are then transported to the cell surface by membrane-bound 'vesicles'. In the case of proteins from invading bacteria, the lipopolysaccharide (LPS) secreted by Gram-negative bacteria is recognized by TLR4 molecules on dendritic cells, which then aids the formation and trafficking of the antigen-presenting complex. Once on the cell surface, the antigen can be recognized by circulating T lymphocyte cells, which initiate a full immune response against the alien molecule. So does a dendritic cell that has eaten particles containing LPS indiscriminately present antigens from all of its phagolysosomes? Blander and Medzhitov² found that phagosomes in the same cell process antigens independently of one another. Those that contain LPS along with the ingested proteins send antigens to the cell surface (Ag2) and those lacking LPS do not (Ag1).

cells contained lipopolysaccharide in other phagosomes. Antigen could be successfully processed from phagosomes containing apoptotic cells if those cells were given lipopolysaccharide before their ingestion. This indicated that activation of TLR4 in the phagosome resulted in selective loading of antigen from that phagosome.

To address possible confounding effects of comparing different kinds of cells as phagocytosed particles, the authors performed analogous experiments using protein antigens attached to synthetic microspheres, with or without added lipopolysaccharide. Antigens were fully processed and presented when the microspheres they were delivered on also contained lipopolysaccharide. By contrast, antigens associated with lipopolysaccharide-free particles did not reach the cell surface, even in cells that had lipopolysaccharide-labelled particles in other phagosomes. Characterization of isolated microparticle-containing phagosomes showed that all of them, with or without lipopolysaccharide, had matured into phagolysosomes. The essential difference was that the phagosomes with lipopolysaccharide particles contained stable MHC class II molecules and showed indications of the advanced stages of antigen processing (that is, proteolysed invariant chain proteins).

Therefore, the organizational unit for this information processing is not the entire cell but rather the individual organelle. Such phagosome autonomy was indicated by a study of antigen presentation by MHC class I proteins from phagosomes⁴, although the independence of the phagosomes was not established as thoroughly as in Blander and Medzhitov's work². The localized selection of antigens for presentation implies that TLR4 signals are generated inside the phagosome, at least with respect to directing the maturation of MHC class II complexes. It also suggests that TLR4 can direct traffic of the membrane organelles that transport protein cargoes from phagosomes to the cell surface. TLR4 signalling was reported to inhibit the trafficking that leads to formation of phagolysosomes⁵, although this point remains controversial^{6,7}.

It is somewhat surprising that the contents of phagosomes remain isolated from each other inside the cells used by Blander and Medzhitov, as studies of other phagocytes have indicated that phagosomes and endocytic organelles readily merge together or exchange contents^{8,9}. Perhaps TLR4 or other lipopolysaccharide-associated activities isolate lipopolysaccharide-rich phagosomes from other kinds of phagosomes.

Several questions emerge from this work, the answers to which could guide the understanding of immune recognition and the design of vaccines. Is the organizational unit smaller than one phagosome? Perhaps patches of phagosomal membranes are sufficient to organize TLR4-dependent MHC class II maturation and trafficking. How are

TLR4-mediated signals that stimulate dendritic-cell maturation integrated with the local signals that drive antigen processing and presentation? Vaccines elicit immune responses using combinations of antigen, particles and TLR ligands, so might vaccine formulations be improved by ensuring the persistent association of these ingredients inside phagosomes? Finally, are there counteracting but as yet undetected signals coming from the apparently inert phagosomes? Maybe every nibbled bit affects the quality of the meal.

Joel Swanson is in the Department of Microbiology and Immunology, University of Michigan Medical School,

Ann Arbor, Michigan 48109-0620, USA.

e-mail: jswan@umich.edu

1. Banchereau, J. & Steinman, R. M. *Nature* **392**, 245–252 (1998).
2. Blander, J. M. & Medzhitov, R. *Nature* **440**, 808–812 (2006).
3. Mellman, I. & Steinman, R. M. *Cell* **106**, 255–258 (2001).
4. Guereamonprez, P. et al. *Nature* **425**, 397–402 (2003).
5. Shiratsuchi, A. et al. *J. Immunol.* **172**, 2039–2047 (2004).
6. Blander, J. M. & Medzhitov, R. *Science* **304**, 1014–1018 (2004).
7. Yates, R. M. & Russell, D. G. *Immunity* **23**, 409–417 (2005).
8. Knapp, P. E. & Swanson, J. A. *J. Cell Sci.* **95**, 433–439 (1990).
9. Berthiaume, E. P., Medina, C. & Swanson, J. A. *J. Cell Biol.* **129**, 989–998 (1995).

SOLAR SYSTEM

When the dust unsettles

Gary R. Huss

Two attempts to measure the isotopic composition of oxygen in the Sun from particles trapped in lunar soils give very different results. A rethink of why the Solar System is as it is might be required.

On page 776 of this issue, Ireland *et al.*¹ report investigations of lunar soil from which they infer that, compared with other Solar System bodies, the Sun is depleted in the naturally most plentiful oxygen isotope, ¹⁶O. Just a year ago, another study² of soils on the Moon concluded exactly the reverse. So who is right?

Oxygen is the third most abundant element in the Solar System (the Sun's huge stores of hydrogen and helium claim first and second place), and a principal constituent of the rocks, ices and atmospheres that make up the planets. Its three naturally occurring isotopes — ¹⁶O, ¹⁷O and ¹⁸O — are found in relative abundances of about 2,700:1:5, although the physical and chemical processes occurring in different environments of the Solar System can cause shifts in these abundances of up to a few per cent.

Such deviations can be depicted on an oxygen three-isotope plot (Fig. 1, overleaf). Here, all samples from Earth fall along a single line with a slope of about 0.5, indicating that the ratio ¹⁸O/¹⁶O shifts by twice as much as the ratio ¹⁷O/¹⁶O. This is mass-dependent behaviour, as the difference in mass between ¹⁸O and ¹⁶O is twice that between ¹⁷O and ¹⁶O. Oxygen from other planets or asteroids — represented on Earth by different classes of meteorite — show similar mass-dependent fractionations, but as the underlying oxygen composition of each body is different, the data fall on different lines on the three-isotope plot. Oxygen isotopic composition has therefore become a crucial parameter in classifying meteorites.

How this oxygen-isotope variability arose, however, is not understood. Does it represent a heterogeneity inherited from the raw

materials that made up the Solar System? Or is it the result of physical or chemical processes in the early Solar System^{3–5}? The initial oxygen isotopic composition of the dust and gas from which our Solar System formed is not known. The Sun contains most of the matter in the Solar System, so its oxygen isotopic composition effectively defines the Solar System's oxygen composition. Models that generate the compositions of other bodies from an ¹⁶O-rich composition are very different from those that start with an ¹⁶O-poor composition. But measuring the chemical and isotopic composition of the Sun directly is difficult; the best data come from measurements of the stream of charged particles emitted by the Sun known as the solar wind.

One of the best places to measure the solar wind is in lunar soils. The Moon has no atmosphere and no magnetic field, so solar-wind particles are implanted directly into its surface. Measurements of matter from the solar wind have so far concentrated on elements such as the noble gases, nitrogen and carbon, which are not themselves significant components of minerals in lunar soil. Oxygen, by contrast, is found in many minerals. Therefore, solar-wind oxygen in lunar soils must be studied in minerals such as iron metal, which are by nature oxygen-free.

Hashizume and Chaussidon² were the first to separate iron-metal particles from lunar soils and measure their oxygen content using an ion microprobe. They found that oxygen in a few grains of their soil, which had been exposed to the solar wind between one billion and two billion years ago, was enriched in ¹⁶O compared with the oxygen on Earth and that

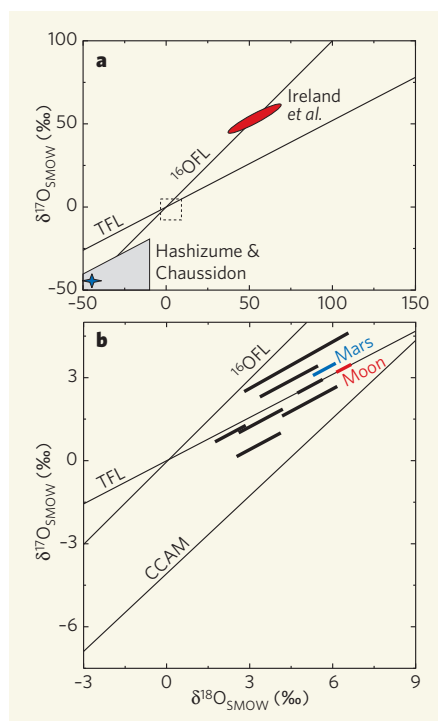
from most other Solar System bodies (Fig. 1). This implied that the Sun itself is ^{16}O -rich, just like the calcium–aluminium-rich inclusions (CAIs) that are found in meteorites and are believed to be among the oldest solid bodies in the Solar System. Now, Ireland *et al.*¹ report results from a contemporary lunar soil that support the opposite conclusion.

The main difference between the two studies is the choice of sample. Hashizume and Chaussidon² used a sample of ancient lunar regolith that they had previously shown contained carbon enriched in ^{13}C (ref. 6) and nitrogen depleted in ^{15}N (ref. 7), results they ascribed to the effect of the solar wind. Their metal grains contained a thick oxide layer that compromised the normal solar-wind implantation profile, but the authors identified ^{16}O -rich oxygen they found inside the grains as so-called solar energetic particles. These particles travel at much higher speeds than normal solar-wind particles, and are therefore much more deeply embedded in the grains.

Ireland and colleagues¹ used a sample of lunar soil only recently exposed to the solar wind, but which had one of the highest exposures known. Their metal grains had only a very thin oxide layer, and the ^{16}O -poor oxygen was found at a depth consistent with normal implanted solar wind. The two studies^{1,2} thus measured solar particles of different ages and, apparently, from different energy regimes. Yet even taking these facts into account, such divergent results are hard to understand.

Processes within the Sun, and those that accelerate solar-wind and solar energetic particles away from its surface, would be expected to affect isotopic composition in a mass-dependent manner. Thus, samples would simply move up or down a line of slope 0.5 on the three-isotope plot. To move away from such a line, either oxygen of a different composition must be added or a process that is mass independent must be invoked. Although the outer layer of the Sun could have changed through the addition of new material over two billion years, there is no known reservoir of material in the Solar System that has a composition extreme enough or a mass great enough to shift the composition of this layer by the required amount.

A process called self-shielding can also alter oxygen isotopic composition. When certain compounds such as carbon monoxide are exposed to intense ultraviolet radiation, molecules containing the highly abundant ^{16}O can exhaust the supply of photons with the correct energy to break them up, although there are still plenty of photons that can disrupt molecules containing ^{17}O and ^{18}O . A reservoir of oxygen with a composition different from the starting carbon monoxide can be created if the released oxygen becomes trapped in a different molecule, such as water⁵. But there is no obvious way to apply this mechanism to solar-wind or solar energetic particles.



We are therefore left with an intriguing dilemma. The data both of Hashizume and Chaussidon² and of Ireland and colleagues¹ are good. Their results cannot, however, be explained within our current understanding of oxygen isotopes and the structure of the Sun. More data should help: a major goal of NASA's Genesis mission, launched in 2001 to capture the solar wind, was to determine the oxygen isotopic composition of the Sun. But the difficulty inherent in measuring oxygen in the Genesis collectors, and the complications introduced when the spacecraft crashed into the Utah desert on its return to Earth in 2004, mean that it will be some time before we have further data.

Figure 1 | Oxygen in the Solar System.

The quantities $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$ give deviations in the ratios $^{17}\text{O}/^{16}\text{O}$ and $^{18}\text{O}/^{16}\text{O}$ relative to the same ratios in standard mean ocean water (SMOW), with $\delta^{17}\text{O} = [(^{17}\text{O}/^{16}\text{O})_{\text{sample}} / (^{17}\text{O}/^{16}\text{O})_{\text{SMOW}} - 1] \times 1,000$, and similarly for $\delta^{18}\text{O}$. **a**, Solar System material lies within the small box at the intersection of two lines: the terrestrial mass fractionation line, TFL, with slope 0.5 and along which all samples from Earth lie; and the ^{16}O fractionation line, ^{16}OFL , with gradient 1. Samples with enhanced or depleted ^{16}O content compared with the SMOW value lie farther up or down the ^{16}OFL line, respectively. The star represents the ratios found in Ca–Al-rich inclusions (CAIs) from meteorites. Hashizume and Chaussidon's results from lunar soils² lie within the grey area; Ireland and colleagues' new findings¹ are in the red disc. **b**, An expansion of the intersect area, showing the lines on which the oxygen isotopic compositions for samples from Mars, the Moon and various meteorite parent bodies sit. The line CCAM refers to an array of measurements of minerals in CAIs.

Depending on the results, some new ideas might be required.

Gary R. Huss is at the Hawaii Institute of Geophysics and Planetology, University of Hawaii at Manoa, 1680 East-West Road, Honolulu, Hawaii 96822, USA.

e-mail: ghuss@higp.hawaii.edu

1. Ireland, T. R., Holden, P., Norman, M. D. & Clark, J. *Nature* **440**, 776–778 (2006).
2. Hashizume, K. & Chaussidon, M. *Nature* **434**, 619–622 (2005).
3. Clayton, R. N., Grossman, L. & Mayeda, T. K. *Science* **182**, 485–488 (1973).
4. Thieme, M. H. & Heidenreich, J. E. III *Science* **219**, 1073–1075 (1983).
5. Clayton, R. N. *Nature* **415**, 460–461 (2002).
6. Hashizume, K. *et al.* *Astrophys. J.* **600**, 480–484 (2004).
7. Hashizume, K., Marty, B. & Wieler, R. *Earth Planet. Sci. Lett.* **202**, 201–216 (2002).

ARCHAEOLOGY

Failure and how to avoid it

Kathleen D. Morrison

Nothing lasts for ever, not least human civilizations. There are many reasons why societies stand or fall, and these lessons from the past require investigation at various places and on various timescales.

Archaeology, it seems, really is a matter of life and death — this was the theme to emerge from a meeting* convened to address the question of what makes societies more likely to collapse or to achieve long-term sustainability. Just as we do today, our ancestors faced problems of resource depletion, environmental degradation, political instability, demographic

pressure and social upheaval. And, as today, success in dealing with these challenges was never assured.

Consider the following contrasts. The islands of eastern Polynesia were all settled within a few centuries of one another by people sharing the same ancestral culture. Yet whereas some islands, such as Tahiti, have sustained human populations for centuries, others, such as Easter Island (Rapa Nui), supported populous and complex societies for

*How Societies Chose to Fail or Succeed, Global Institute of Sustainability Workshop, Arizona State University, Tempe, Arizona, USA, 31 January to 2 February 2006.

only a short time before experiencing profound demographic and social disruption.

Completely isolated since its initial colonization, variously dated between about AD 750 and 1200^{1,2}, this scrap of land came to the notice of the world with the visit of the Dutch explorer Jacob Roggeveen on Easter Sunday, 1722. Roggeveen marvelled not only at the more than 200 massive stone statues, the Moai, which ring the coast, but also at the barrenness of the landscape and the destitution of its small population. The rich soils of Easter Island, oddly enough, supported a depauperate vegetation virtually devoid of woody plants.

That Easter Island had once supported a much larger population with a complex political structure was clear from the archaeological record. But it took a combination of archaeological and palaeoenvironmental data (pollen, microscopic charcoal and faunal analysis) to reveal the extent to which the island's degraded landscape was a product of human action. Once covered by subtropical forests dominated by a now-extinct species of large palm^{3,4}, the island environment was ravaged by intensive human exploitation. With the destruction of plants for canoe-building, off-shore food resources such as the marine mammals exploited early in the island's human history receded from reach, as did any chance of mobility as an option for addressing the mismatch between resources and needs.

Although ecological constraints certainly played a role in the variable success of human populations on Pacific islands (B. Rolett, Univ. Hawaii), cultural practices were clearly also crucial. Longer-term sustainability involved agricultural systems and levels of resource extraction compatible with local conditions.

Not all anthropogenic environmental change has led to cultural collapse, as both the contrasts between Pacific islands and the evidence from more complex continental contexts shows. One such comparison is between the long-term occupation of the Basin of Mexico and the well-studied collapse of the Classic-period Maya (Fig. 1). The Maya, a complex urban society organized into a series of competitive city states, abruptly ceased building monumental structures around the ninth century AD. Large parts of the Maya homeland were depopulated, although others continued to thrive. Many factors — including environmental degradation, population expansion, warfare and a decline in the ideology of kingship — seem to have been at issue in the collapse (D. Webster, Penn. State Univ.). Not far away, however, in central Mexico, the city of Teotihuacan managed to maintain a large population, perhaps as many as 100,000 people, for more than 400 years (G. Cowgill, Arizona State Univ.). The area has continued to support a large population to this day.

The experiences of past societies may be thought of as social and ecological 'experiments' (C. Redman, Arizona State Univ.). In each case, past conditions, responses to change



Figure 1 | Societal collapse: the case of the Classic Maya. The complex urban way of life of the Maya ended abruptly between the ninth and tenth centuries AD. Monumental construction for the élite (such as Temple 1 at Tikal, left) ceased entirely. Political collapse was accompanied by large-scale depopulation, indicating problems at all levels of society. Archaeological and palaeoenvironmental data indicate environmental degradation caused by intensive agriculture (upper right, raised and ditched fields from the Upper Booth's River in Belize). In contrast, the great city of Teotihuacan in central Mexico (lower right, Pyramid of the Sun and the Avenue of the Dead in the city centre) lasted from about 100 BC to AD 650, with continued intensive occupation of the region to the present. In each case, both social strategies and environmental factors influenced the chances of long-term success.

(real or perceived) and outcomes of behaviour constitute data for broader comparisons. Defining cases requires more precise definitions, drawing distinctions between restorations of continuity, societal transformations, failures or 'mini-collapses' such as dynastic transitions, and collapse proper (Redman). Considerations of spatial and temporal scale are essential, not least in defining sustainability. The historical record represents the cumulative outcome of multiple processes (S. van der Leeuw, Arizona State Univ.), so explanations must integrate processes and contingent events operating at different scales.

Several factors seem to have determined whether societies failed, collapsed, or experienced either episodic change or 'radical continuity'. On the one hand, environmental changes, natural or self-inflicted, caused or assisted many cases of collapse and failure (J. Betancourt, Univ. Arizona). Other factors include rates of demographic change (Webster), the degree to which élites were isolated (and insulated) from social problems, the strength of communication across hierarchies (S. Schroeder, Univ. Wisconsin), levels of investment in infrastructure, and the ability to balance commitment to cultural values with the flexibility required to manage uncertainty (J. Diamond, UCLA; M. Nelson, K. Spielmann, Arizona State Univ.).

In the past, many societies with fewer vulnerabilities than today's Western societies have failed, whereas others with more vulnerabilities have succeeded (B. Nelson, Arizona State Univ.). The long archaeological record of human history constitutes the best and, for most of our history, the only source of data about the long-term consequences of human choices. One of the values of archaeological research is that it offers a rich source of data on human actions and their consequences — data that may someday play the same kind of role in understanding present challenges that palaeoecology now does in climate modelling. By identifying the differences, as well as the similarities, between past and present societies, archaeology can publicize environmental and social risks and vulnerabilities while underscoring an increasingly wide range of technical and cultural solutions (Betancourt).

What lessons does current archaeological research have for policy-makers? For one, it is clear that large-scale environmental degradation is almost always a factor in social collapse. Awareness of ecological inflexion points, beyond which recovery is no longer possible, may make the difference between success and failure. Total collapse of a civilization may be rare, but cultural transformations are not only common but perhaps also necessary. For example, the long-term occupation of parts of

south Asia may have been made possible by a willingness to experiment with new social forms and technological strategies, including radical changes in agriculture and cuisine. In today's globally connected world, failures tend to have ramifications well beyond local contexts, making the lessons of the past perhaps even more relevant. ■

Kathleen D. Morrison is in the Department of Anthropology, University of Chicago, 1126 East 59th Street, Chicago, Illinois 60637, USA. e-mail: morrison@uchicago.edu

1. Green, R. C. & Weisler, M. I. *Asian Perspect.* **41**, 213–241 (2002).
2. Hunt, T. L. & Lipo, C. P. *Science* **311**, 1603–1606 (2006).
3. Flenley, J. R. & King, S. M. *Nature* **307**, 47–50 (1984).
4. Dransfield, J. et al. *Nature* **312**, 750–752 (1984).

FLUID DYNAMICS

The rough with the smooth

Kwing-So Choi

Those who go with the flow assert that rough surfaces cause turbulence in fluids passing over them. The claim that, under certain conditions, the opposite is possible disturbs that cherished belief.

Have you noticed how slick the exteriors of cars have become lately? Side mirrors are aerodynamically shaped; windows are an integral part of the doors; no parts hang out over the side of the body. Even windscreen wipers can be retracted when not in use. A smooth body is not only better looking; it can reduce aerodynamic drag, and thus makes a car faster and more fuel-efficient — or so we are led to believe. In fact, this prevailing wisdom may not be true: Fransson *et al.*¹ report in *Physical Review Letters* that surface bumps are actually useful for reducing drag. So what's going on?

To get to the bottom of this matter, we must look at the evolution of fluid flows over a solid surface. Initially, these flows are silky smooth and move together; they are 'laminar'. Soon, however, the inherent instability of a laminar flow leads to its becoming random and chaotic, or 'turbulent'. Turbulent flow is usually associated with characteristics such as strong mixing of the fluid and high energy loss through dissipation. The drag that cars experience is therefore greater when the surrounding flow is turbulent: if a laminar flow can be maintained, the frictional drag acting on the car is much less.

Fransson and colleagues¹ conducted experiments under controlled conditions in a wind tunnel, and were able to maintain laminar flow along a wall for longer by deliberately making its surface rough. We naively expect that surface roughness will create disturbances that lead to turbulent flow, so the result is, on the face of it, puzzling. But the authors were able to show that, by carefully controlling the disturbances created by the surface roughness, they could make the flow over a body surface more resistant to a certain type of instability, and so postpone the onset of turbulence until farther downstream (Fig. 1).

The authors used a flat wall, to which uniformly distributed 'roughness' elements were attached that produced steady, corkscrew-like disturbances in the flow, called streaks. Their experimental results confirm simulations showing² that a moderate strength of streaks can indeed stabilize the laminar flow. Although the size and shape of roughness elements are important in this experiment, the strength of the disturbance they create in the flow is also crucial. If this is too strong, a more deadly instability kicks in and the flow quickly becomes turbulent³.

Using surface roughness to reduce friction drag is, in the case of turbulent flow, not new. Randomized, V-shaped roughness elements are known to reduce friction drag⁴ in such flows, as are sand grains attached to a surface⁵. Cross-flow variation — a disturbance in velocity that occurs at right angles to the main flow — is also known to be caused by regular roughness elements. Its effect on turbulent flow has been studied theoretically⁶ by solving the equations governing fluid flows directly by computer.

In these cases, the key to drag reduction is efficient modification of organized structures of various shapes that have been identified as supplying most of the energy in turbulent flows. It has been argued that randomized roughness elements destroy such structures⁴, whereas cross-flow variation stabilizes a flow⁶, thus preventing the structures' reproduction. A regular pattern of roughness elements that introduces such a cross-flow variation to a laminar flow is central to Fransson and colleagues' study¹. Their success in reducing drag reinforces the idea that organized structures in turbulent flows are the remnants of laminar flow structures after they have undergone the transition to turbulence.

Several investigations have sought to repeat physical and numerical experiments by artificially introducing disturbances into the flow, but only a few⁷ have succeeded. This implies that such results might be sensitive to small variations in geometric and flow parameters. So how robust is the flow-control strategy proposed by Fransson and colleagues?

It is fair to say that, as the shape of the roughness elements is critically important in stabilizing laminar flow, totally different flow patterns might emerge even when their shape is changed only slightly. Applying the technique to high-speed flows might also present a problem, as the streaks will become unsteady at higher speed. Roughness elements will introduce additional drag as they disturb the flow, so any reduction in friction drag by roughness elements must compensate for this if a strategy is to have practical value.

The prize for overcoming such problems, and recognizing that smooth is not everything, is great. Its value lies in the ability to design and build fuel-efficient airplanes, quieter submarines and faster cars — both more easily and more cheaply. ■

Kwing-So Choi is in the School of Mechanical, Materials and Manufacturing Engineering, University of Nottingham, University Park, Nottingham NG7 2RD, UK.

1. Fransson, J. H. M., Talamelli, A., Brandt, L. & Cossu, C. *Phys. Rev. Lett.* **96**, 064501 (2006).
2. Cossu, C. & Brandt, L. *Phys. Fluids* **14**, L57–L60 (2002).
3. Andersson, P. et al. *J. Fluid Mech.* **428**, 29–60 (2001).
4. Sirovich, L. & Karisson, S. *Nature* **388**, 753–755 (1997).
5. Tani, I. *Proc. Japan Acad. B* **64**, 21–24 (1988).
6. Schoppa, W. & Hussain, F. *Phys. Fluids* **10**, 1049–1051 (1998).
7. Iuso, G. et al. *J. Fluid Mech.* **473**, 23–58 (2002).

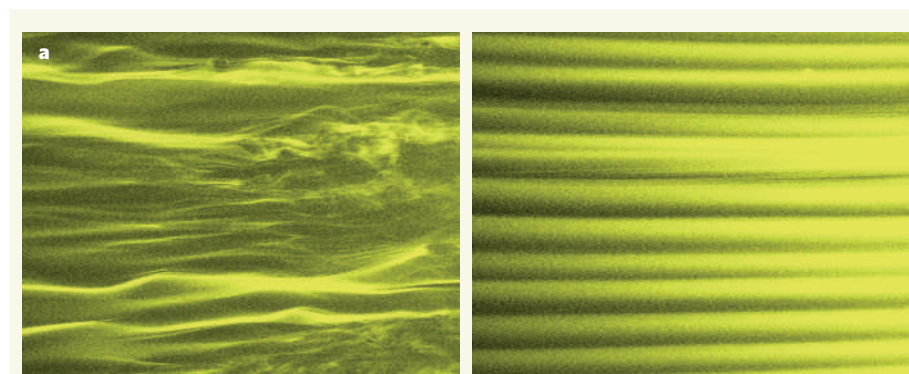


Figure 1 | Flow patterns visualized by smoke. **a**, Laminar flow goes through the transition to turbulence without surface roughness. **b**, Fransson *et al.*¹ maintain laminar flow for longer with a rough surface that creates regular disturbances.

BRIEF COMMUNICATIONS

Early Neolithic tradition of dentistry

Flint tips were surprisingly effective for drilling tooth enamel in a prehistoric population.

Prehistoric evidence for the drilling of human teeth *in vivo* has so far been limited to isolated cases from less than six millennia ago^{1–3}. Here we describe eleven drilled molar crowns from nine adults discovered in a Neolithic graveyard in Pakistan that dates from 7,500–9,000 years ago. These findings provide evidence for a long tradition of a type of proto-dentistry in an early farming culture.

The site of Mehrgarh in Baluchistan lies along the principal route connecting Afghanistan to the Indus valley. After intermittent occupations by hunter-gatherers, Mehrgarh's subsistence economy shifted to the cultivation of barley and wheat, cotton domestication and cattle breeding⁴. Diachronic archaeological evidence records an increasingly rich cultural life, with technological sophistication based on diverse raw materials. Excavation of the Neolithic cemetery known as MR3 yielded more than 300 graves created over a 1,500-year time span⁴.

We identified four females, two males and three individuals of unknown gender that between them had a total of eleven drilled permanent crowns (for details, see supplementary information). Drilled teeth were evident in both jaws (four in the maxilla; seven in the mandible) and exclusively in the first (four specimens) or second (seven specimens) permanent molars. Except for one located on the distal-buccal cervix of a lower first molar, the holes were bored into enamel or secondary dentine on the occlusal surfaces. The holes are 1.3–3.2 mm in diameter and angled slightly to the occlusal plane, with a depth of between 0.5 and 3.5 mm.

Light microscopy, scanning electron microscopy and microtomography revealed cavity shapes that were conical, cylindrical or trapezoidal (Fig. 1; and see supplementary information). They also showed concentric ridges preserved on some walls that had been left by the drilling tool. The teeth of at least one individual reveal that the procedure involved not just removal of the tooth structure by the drill, but also subsequent micro-tool carving of the cavity wall by either the operator or the patient. In all cases, marginal smoothing confirms that drilling was performed on a living person who continued to chew on the tooth surfaces after they had been drilled.

Complete dentition for the 11 specimens is not preserved, so the incidence per jaw cannot be determined. However, one individual had

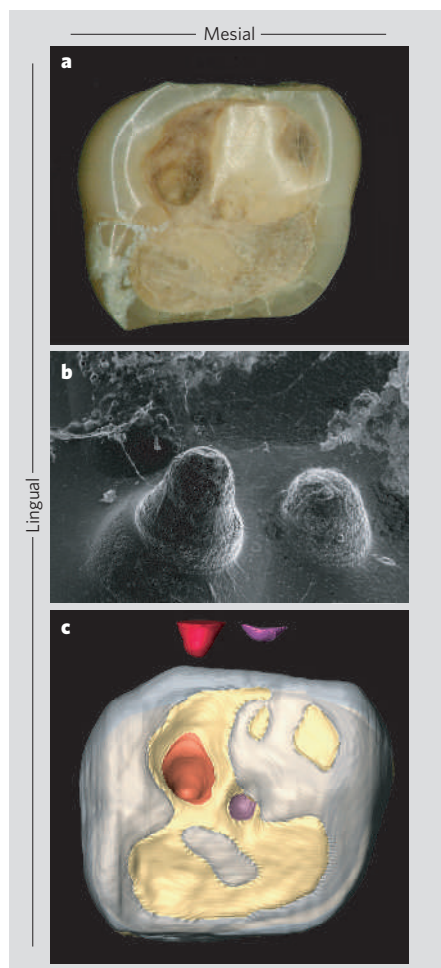


Figure 1 | Maxillary left second molar from an adult male (MR3 90) from Neolithic Mehrgarh.

There are two *in vivo* perforations on the occlusal surface made by a drilling tool that was probably equipped with the same flint head. **a**, The larger, mesio-lingual perforation has a maximum diameter of 1.6 mm; the second, at the centre of the crown, has a maximum diameter of 1.3 mm. **b**, Scanning electron micrograph of their negative replicas, showing that both perforations are slightly inclined mesio-distally and have a similar general shape. The larger perforation is also deeper (1.5 compared with 0.7 mm). **c**, A microtomographic three-dimensional reconstruction of the tooth, with positive virtual casts (top) of the two perforations (for methods, see supplementary information). The minimum volume of removed enamel and dentine in the mesio-lingual perforation (red) is 1.8 mm³; the minimum volume of removed enamel for the smaller perforation (violet) is 0.3 mm³. Scale bars: **a**, 2.2 mm; **b**, 1 mm; **c**, 2 mm.

three drilled teeth and another had a tooth that had been drilled twice. Four teeth show signs of decay associated with the hole, indicating that the intervention in some cases could have been therapeutic or palliative. But as caries can exist in individuals in the absence of drilling, and drilling may be done in individuals without caries, the motive for the early Neolithic dental procedure described here is unclear. Aesthetic functions⁵ can be ruled out because of the deep placement in the jaw. The perforations exposed sensitive tooth structure, so some type of filling may have been placed in the cavity; however, no evidence survives to confirm this.

Whatever the purpose, tooth drilling on individuals buried at MR3 continued for about 1,500 years, indicating that dental manipulation was a persistent custom. After 6,500 yr BP, the practice must have ceased, as there is no evidence of tooth drilling from the subsequent MR2 Chalcolithic cemetery, despite the continuation of poor dental health⁶.

At Neolithic Mehrgarh, flint drill heads occur in the lithic assemblages associated with beads of bone, steatite, shell, calcite, turquoise, lapis lazuli and carnelian⁷. Using models of these drill tips, we reconstructed a method for drilling based on the ethnographic literature⁸ and found that a bow-drill tipped with a flint head required less than one minute to produce similar holes in human enamel. Presumably, the know-how originally developed by skilled artisans for bead production was successfully transferred to drilling of teeth in a form of proto-dentistry.

A. Coppa*, L. Bondioli†, A. Cucina‡, D. W. Frayer§, C. Jarrige||, J.-F. Jarrige||, G. Quivron||, M. Rossi¶, M. Vidale#, R. Macchiarelli☆

*Dipartimento BAU, Università 'La Sapienza', 00185 Rome, Italy

†Sezione di Antropologia, Museo Nazionale Preistorico Etnografico 'L. Pigorini', 00144 Rome, Italy

‡Facultad de Ciencias Antropologicas, Universidad Autonoma de Yucatan, 455 LL 41-43 Centro Merida, Yucatan, Mexico

§Department of Anthropology, University of Kansas, Lawrence 66045-2110, USA

||Musée National des Arts Asiatiques Guimet, 75116 Paris, France

¶Dipartimento di Fisica, Università di Bologna, 40127 Bologna, Italy

#Istituto Centrale per il Restauro, 00184 Rome, Italy

☆Laboratoire GBPH, UMR 6046, Université de Poitiers, 86022 Poitiers, France
e-mail: roberto.macchiarelli@univ-poitiers.fr

1. Bennike, P. & Fredebo, L. *Bull. Hist. Dent.* **34**, 81–87 (1986).
2. Schwartz, J. H., Brauer, J. & Gordon-Larsen, P. *Am. J. Phys. Anthropol.* **97**, 77–85 (1995).
3. White, T. D., Degusta, D., Richards, G. D. & Baker, S. G. *Am. J. Phys. Anthropol.* **103**, 409–414 (1997).
4. Jarrige, C., Jarrige, J.-F., Meadow, R. H. & Quivron, G. (eds) *Mehrgarh Field Reports 1974–1985. From Neolithic Times to Indus Civilization* (Department of Culture & Tourism, Karachi, 1995).
5. Alt, K. W. & Pichler, S. L. in *Dental Anthropology* (eds Alt, K. W., Rosing, F. W. & Teschler-Nicola, M.)

389–415 (Springer, Vienna, 1998).

6. Lukacs, J. R. & Mindermann, L. L. in *South Asian Archaeology 1989* (ed. Jarrige, C.) 167–179 (Prehistory Press Monographic Publication, Madison, 1992).
7. Kenoyer, J. M. & Vidale, M. in *Material Issues in Art and Archaeology III* (eds Vandiver, P. B., Druzik, J. R., Wheeler, G. S. & Freestone, I. C.) 495–518 (Material Research Society, Pittsburgh, 1992).
8. Kenoyer, J. M., Vidale, M. & Bhan, K. K. *World Archaeol.* **23**, 44–63 (1991).

Supplementary information accompanies this communication on Nature's website.

Received 22 November 2005; accepted 1 March 2006.

Competing financial interests: declared none.

doi:10.1038/440775a

PARASITOLOGY

Parasite survives predation on its host

As prisoners in their living habitat, parasites should be vulnerable to destruction by the predators of their hosts. But we show here that the parasitic gordian worm *Paragordius tricuspidatus* is able to escape not only from its insect host after ingestion by a fish or frog but also from the digestive tract of the predator. This remarkable tactic enables the worm to continue its life cycle.

The induced suicide of crickets infected by gordian worms is one of the best known examples of parasite manipulation of host behaviour¹. Adult gordian worms are free-living in water, where they mate as a knotted mass of multiple individuals. Emergence from the host occurs only after the cricket enters

the water (for movie, see supplementary information) and may take as long as 10 min owing to the large size of the worm¹. During this time, the cricket is active at the surface and attractive to aquatic predators such as fish and frogs (F.T., unpublished observations). Death of the worm would be expected to result from generalist predation upon the host at this stage unless the parasite were capable of an antipredator response.

Few parasites have their own predators, although they are victims of those of their hosts. Predation upon a host may shape parasite life history in two important ways. First, increased predation may select for increased parasite virulence^{2,3}: for example, parasite

development can speed up to minimize the period of host occupancy⁴. Increased virulence may also take the form of manipulating the host's behaviour⁵, for example to remove it from sources of predation^{6,7}. Second, predation may affect parasite life history if the predator becomes incorporated into the life cycle^{8–10}.

We investigated the response of gordian worms to predation on their host. Under laboratory conditions, we found that crickets that harboured or were expelling gordian worms were often eaten by generalist predators — fish (trout (*Oncorhynchus mykiss*), perch (*Lepomis gibbosus*), bass (*Micropterus salmoides*)) and frogs (*Rana erythraea*). In none of the 477 predation events that we observed did the predator regurgitate the cricket. Remarkably, the worm escaped predation by wriggling out of the mouth, nose or gills of the predator that had consumed its host (Fig. 1). This escape was recorded from trout (18%, $n = 141$), bass (26%, $n = 292$), perch (22%, $n = 27$) and frogs (35%, $n = 17$) (for methods, see supplementary information).

The mean time until full emergence was 8.6 min (518 ± 208 s, $n = 72$). The maximum time was 28 min, when the predator repeatedly tried to swallow the worm while it was escaping. If a worm did not start to emerge from the mouth, gills or nose within 5 min, it failed to escape — dying, presumably, in the hostile environment of the predator's stomach. To our knowledge, this escape response by a gordian worm is the first example of a parasite or any organism surviving predation in this way.

Fleur Ponton*, Camille Lebarbenchon*†, Thierry Lefèvre*, David G. Biron*,

David Duneau*, David P. Hughes‡, Frédéric Thomas*

***GEMI, UMR CNRS/IRD 2724, IRD, BP 64501,**

34394 Montpellier cedex 5, France

e-mail: fleur.ponton@mpl.ird.fr

†Station Biologique de la Tour du Valat,

Le Sambuc, 13200 Arles, France

‡Centre for Social Evolution, Institute of Biology,

21000 Copenhagen, Denmark

§Centre for Social Evolution, Institute of Biology,

21000 Copenhagen, Denmark

¶Centre for Social Evolution, Institute of Biology,

21000 Copenhagen, Denmark

1. Thomas, F. et al. *J. Evol. Biol.* **15, 356–361 (2002).**

2. Day, T. *Proc. R. Soc. Lond. B* **269, 1317–1323 (2002).**

3. Ebert, D. & Herre, E. A. *Parasitol. Today* **12, 96–101 (1996).**

4. Harvey, J. A. & Strand, M. R. *Ecology* **83, 2439–2451 (2002).**

5. Read, A. F. *Trends Microbiol.* **2, 73–76 (1994).**

6. Moore, J. *Parasites and the Behavior of Animals* (Oxford Univ. Press, Oxford, 2002).

7. Brodeur, J. *Norw. J. Agr. Sci.* **16, 147–153 (1994).**

8. Choisy, M., Brown, S. P., Lafferty, K. D. & Thomas, F. *Am. Nat.* **162, 172–181 (2003).**

9. Parker, G. A., Chubb, J. C., Ball, M. A. & Roberts, G. N. *Nature* **425, 480–484 (2003).**

10. Lafferty, K. D. *Parasitol. Today* **15, 111–115 (1999).**

Supplementary information accompanies this communication on Nature's website.

Received 3 January; accepted 10 March 2006.

Competing financial interests: declared none.

doi:10.1038/440776a

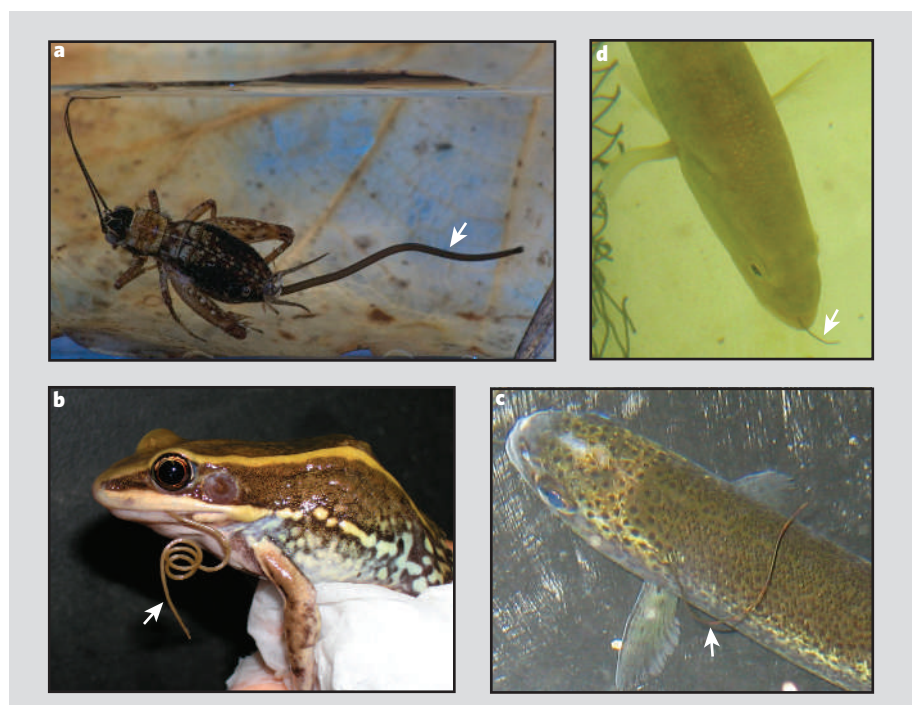


Figure 1 | Escape of parasitic gordian worms from their insect host and from the host's predators. a, Gordian worm (arrow) emerging from a host cricket; b–d, gordian worms (arrows) emerging from a frog (b), a trout (c) and a bass (d) after ingestion of the host insect by these predators. For movies, see supplementary information.

ANIMAL BEHAVIOUR

Chimpanzee choice and prosociality

Arising from: J. B. Silk *et al.* *Nature* **437**, 1357–1359 (2005)

Silk *et al.* report that adult chimpanzees show no difference in their choices in a situation where one choice benefits a familiar conspecific and the other does not¹. From this, they conclude that chimpanzees are indifferent to the welfare of unrelated group members. But without additional data confirming that chimpanzees do choose differently in circumstances in which a difference would be expected, the authors cannot conclude that there is no difference in their scenario. How chimpanzees react to the welfare of unrelated group members remains an open question.

Silk *et al.* evaluate choice behaviour in adult chimpanzees given two options: one in which the actor is provided with food (1/0 option) and another in which both the actor and an unrelated adult chimpanzee in an adjoining enclosure are provided with food (1/1 option). They found no significant differences between the options: actors chose the 1/1 option 56% of the time (in the Louisiana group) or 48% (in the Texas group) when there was no chim-

panzee in the adjoining enclosure and 58% (Louisiana) or 48% (Texas) of the time when there was another chimpanzee in the adjoining enclosure.

In a control experiment in which food was available for only one option, actors chose that option 92% of the time when they were alone and 94% of the time when another chimpanzee was in the adjoining room. This showed that actors were responding to the presence of food in the dish that became available to them.

However, we contest that the authors' conclusion that chimpanzees are indifferent to the welfare of unrelated group members is flawed, because it depends on the null hypothesis. A failure to observe different behaviours in the two different conditions does not prove that those two conditions will always yield the same result. More evidence is needed that chimpanzees choose to benefit a conspecific in conditions in which they would be expected to show such a preference. For example, would

chimpanzees choose the 1/1 option more frequently when a genetically related chimpanzee is in the adjoining chamber? Without supporting information of this type, the authors' premise — that, if chimpanzees are concerned for another's welfare, they should choose the 1/1 option more often when another chimpanzee is present than when they are alone — is unsubstantiated.

Silk *et al.* address a question with implications for the evolution of altruistic behaviour in humans and other primates. This makes it all the more important that their conclusions should be backed up by strong evidence.

Richard J. Beninger, Vernon L. Quinsey
Departments of Psychology and Psychiatry,
Queen's University, Kingston, Ontario K7L 3N6,
Canada
e-mail: beninger@post.queensu.ca

1. Silk, J. B. *et al.* *Nature* **437**, 1357–1359 (2005).

doi:10.1038/nature04758

ANIMAL BEHAVIOUR

Silk *et al.* replyReplying to: R. J. Beninger & V. L. Quinsey *Nature* **440**, doi:10.1038/nature04758 (2006)

Beninger and Quinsey¹ argue that we provide no evidence that chimpanzees show other-regarding preferences in the two-option test situation under conditions in which they would be expected to show such a preference. This criticism is misdirected, because our aim was not to determine whether chimpanzees would demonstrate prosocial preference under any circumstances. Instead, it was to determine whether chimpanzees show prosocial preferences in situations similar to those in which these occur routinely in humans.

Our results², which have been replicated in an independent study of a different group of chimpanzees³, show that they behave very differently from the way we would expect humans to behave. In a study very similar to ours, 3–5-year-old children were asked whether they would prefer to have one sticker for themselves and one sticker for a young

female experimenter, or just one sticker for themselves⁴. Most of the children chose the prosocial option, and some were even willing to give up stickers to the experimenter.

Moreover, we question Beninger and Quinsey's assumption that chimpanzees would behave differently towards kin and non-kin in our experimental protocol. Although female chimpanzees sometimes share plant foods with their infants⁵, analysis of these food transfers indicates that the infants take the initiative in these events, not the mothers. Mothers offer their infants only the unpalatable parts of their own food, and generally seem reluctant to pass them nutritious foods⁶. The role of kinship in cooperation among adults is also uncertain, as cooperation among males often involves reciprocal exchanges between unrelated partners⁷. **Joan B. Silk*, Sarah F. Brosnan†‡, Jennifer Vonk§, Joseph Henrich†, Daniel J. Povinelli||,**

Amanda S. Richardson‡, Susan P. Lambeth‡, Jenny Mascaro†, Steven J. Shapiro‡

*Department of Anthropology, University of California, Los Angeles, California 90095, USA
e-mail: jsilk@anthro.ucla.edu

†Department of Anthropology, Emory University, Atlanta, Georgia 30322, USA

‡Michale E. Keeling Center for Comparative Medicine and Research, University of Texas M. D. Anderson Cancer Center, Texas 78602, USA

§Department of Psychology, University of Southern Mississippi Gulf Coast, Long Beach, Mississippi 39560, USA

||Cognitive Evolution Group, University of Louisiana at Lafayette, New Iberia, Louisiana 70560, USA

1. Beninger, R. J. & Quinsey, V. L. *Nature* **440**, doi:10.1038/nature04758 (2006).
2. Silk, J. B. *et al.* *Nature* **435**, 1357–1359 (2005).
3. Jensen, K., Hare, B., Call, J. & Tomasello, M. *Proc. R. Soc. Lond. B* doi:10.1098/rspb.2005.3417 (2006).
4. Thompson, C., Barressi, J. & Moore, C. *Cogn. Dev.* **12**, 199–212 (1997).
5. Nishida, T. & Turner, L. *Int. J. Primatol.* **17**, 947–968 (1996).
6. Ueno, A. & Matsuzawa, T. *Primates* **45**, 231–239 (2004).
7. Mitani, J., Merriwether, D. A. & Zhang, C. *Anim. Behav.* **59**, 885–893 (2000).

doi:10.1038/nature04759

A Devonian tetrapod-like fish and the evolution of the tetrapod body plan

Edward B. Daeschler¹, Neil H. Shubin² & Farish A. Jenkins Jr³

The relationship of limbed vertebrates (tetrapods) to lobe-finned fish (sarcopterygians) is well established, but the origin of major tetrapod features has remained obscure for lack of fossils that document the sequence of evolutionary changes. Here we report the discovery of a well-preserved species of fossil sarcopterygian fish from the Late Devonian of Arctic Canada that represents an intermediate between fish with fins and tetrapods with limbs, and provides unique insights into how and in what order important tetrapod characters arose. Although the body scales, fin rays, lower jaw and palate are comparable to those in more primitive sarcopterygians, the new species also has a shortened skull roof, a modified ear region, a mobile neck, a functional wrist joint, and other features that presage tetrapod conditions. The morphological features and geological setting of this new animal are suggestive of life in shallow-water, marginal and subaerial habitats.

The evolution of tetrapods from sarcopterygian fish is one of the major transformations in the history of life and involved numerous structural and functional innovations, including new modes of locomotion, respiration and hearing. Fish and tetrapod fossils across this transition can reveal how these innovations were assembled. During the origin of tetrapods in the Late Devonian (385–359 million years ago), the proportions of the skull were remodelled, the series of bones connecting the head and shoulder was lost, and the region that was to become the middle ear was modified. At the same time, robust limbs with digits evolved, the shoulder girdle and pelvis were altered, the ribs expanded, and bony connections between vertebrae developed. Few of these features, however, are seen in the closest relatives of tetrapods—the elpistostegalian fishes—which are incompletely known. *Elpistostege*, for example, is represented only by two partial dermal skull roofs and a segment of the axial skeleton from the early Frasnian Escuminac Formation in Quebec^{1–3}. The best-known elpistostegalian, *Panderichthys*, consists of complete specimens of Middle to Late Devonian age (late Givetian and early Frasnian stages) mostly from the Lode quarry in Latvia^{4–10}. *Panderichthys* possesses relatively few tetrapod synapomorphies, and provides only partial insight into the origin of major features of the skull, limbs and axial skeleton of early tetrapods. In view of the morphological gap between elpistostegalian fish and tetrapods, the phylogenetic framework for the immediate sister group of tetrapods has been incomplete and our understanding of major anatomical transformations at the fish–tetrapod transition has remained limited.

The discovery of a new elpistostegalian sarcopterygian from the Fram Formation in Nunavut Territory, Canada (Fig. 1) significantly enhances our knowledge of the fish–tetrapod transition. Many articulated specimens from a single site are used to describe a taxon that is a remarkable intermediate between *Panderichthys* and early tetrapods. The material provides opportunities to assess the morphological and functional changes associated with the origin of tetrapods.

Geological framework

The Fram Formation is the proximal, continental facies of a Middle–Upper Devonian clastic wedge distributed widely across the

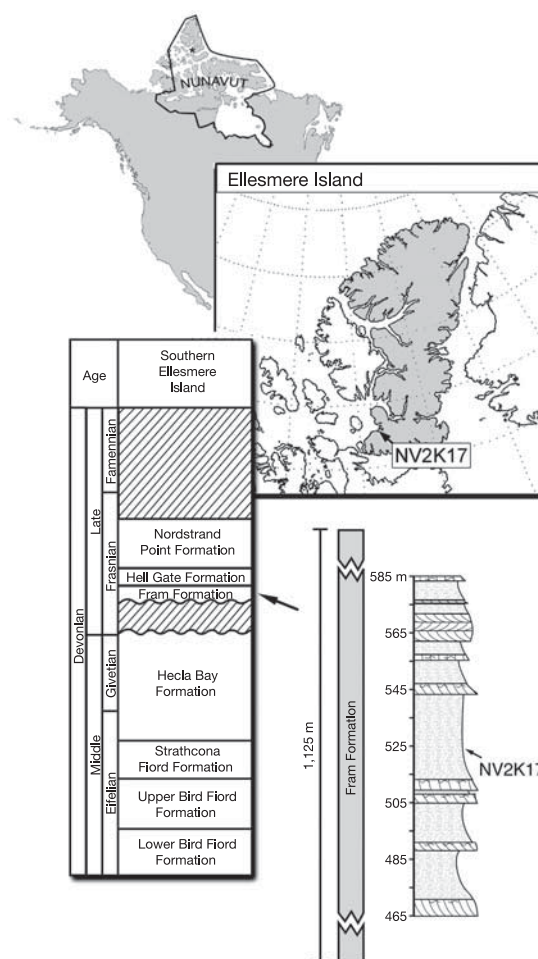


Figure 1 | Geographic location and stratigraphic position of the discovery site (NV2K17) on southern Ellesmere Island, Nunavut Territory, Canada.

¹Academy of Natural Sciences, 1900 Benjamin Franklin Parkway, Philadelphia, Pennsylvania 19103, USA. ²University of Chicago, Department of Organismal Biology and Anatomy, 1027 E. 57th Street, Chicago, Illinois 60637, USA. ³Harvard University, Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, 26 Oxford Street, Cambridge, Massachusetts 02138, USA.

Canadian Arctic Archipelago. Palynomorph biostratigraphy places the Fram Formation in the early and middle Frasnian stage of the Late Devonian^{11–13}. Elpistostegals were found at three localities in the Fram Formation, but the most fossiliferous site (NV2K17) lies in the type section approximately 500 m above the base of the 1,125-m-thick sequence (Fig. 1). The Fram Formation, characterized by alternating resistant sandstones and recessive siltstones, is interpreted as the deposits of meandering stream systems^{11,13}.

Site NV2K17 is within a 30-m-thick, siltstone-dominated sequence bounded by cross-bedded channel sandstones. The fossiliferous unit is a 15-cm-thick, poorly sorted siltstone with dense concentrations of carbonate nodules, intraformational clasts, and skeletal fragments overlain by a 15-cm-thick massive siltstone with articulated and disarticulated remains of sarcopterygian fishes. This package of sediment is evidence of a channel avulsion event that carried bedload, suspended sediment load, and fishes into an inter-channel area where rapid deposition occurred.

Articulated individuals occur as three-dimensional, slightly crushed specimens scattered across the fossiliferous zone, often overlapping and with no preferred orientation. The upper surfaces of articulated specimens are usually more incomplete than the lower,

ferous unit is a 15-cm-thick, poorly sorted siltstone with dense concentrations of carbonate nodules, intraformational clasts, and skeletal fragments overlain by a 15-cm-thick massive siltstone with articulated and disarticulated remains of sarcopterygian fishes. This package of sediment is evidence of a channel avulsion event that carried bedload, suspended sediment load, and fishes into an inter-channel area where rapid deposition occurred.

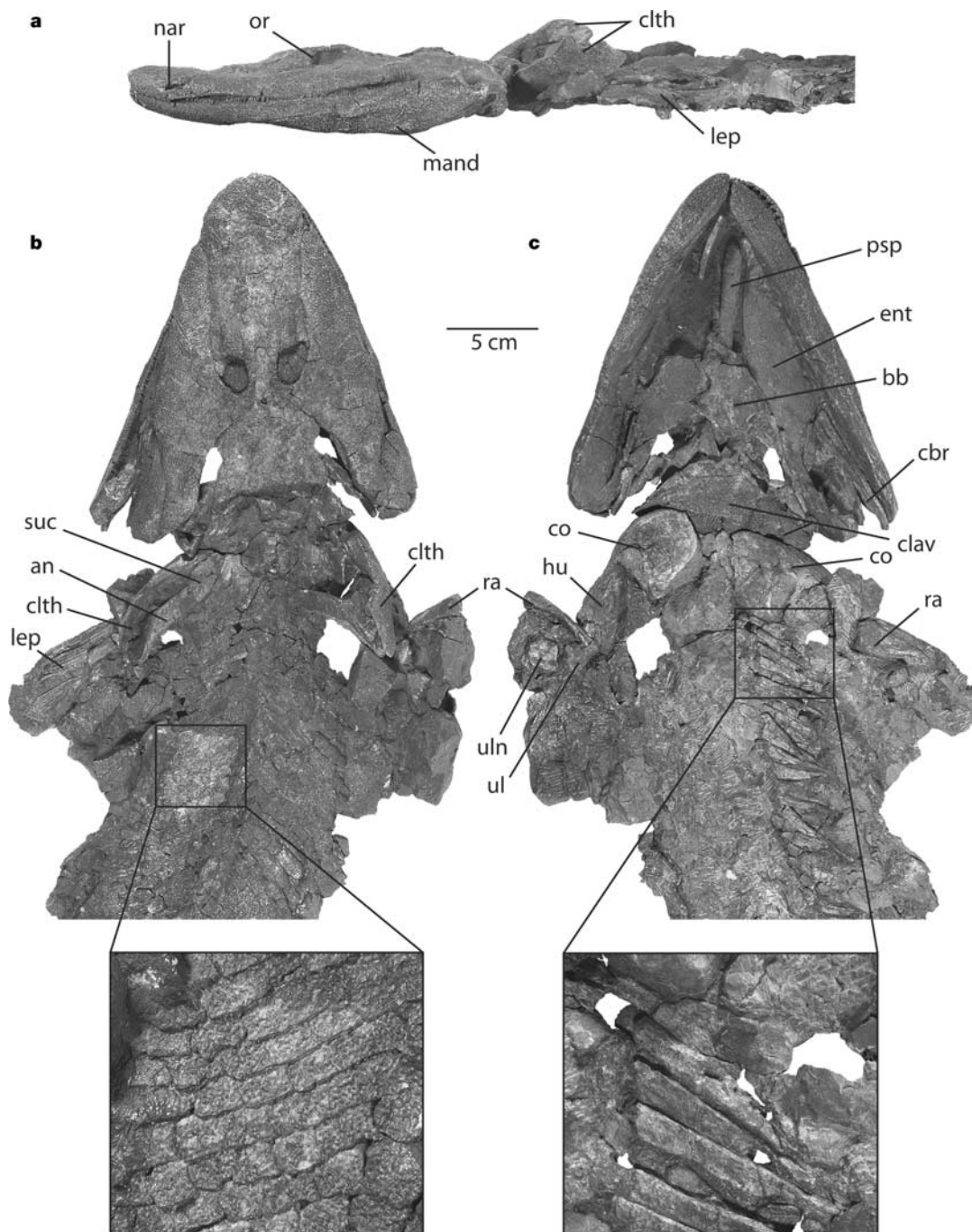


Figure 2 | NUFV 108, holotype of *Tiktaalik roseae* gen. et sp. nov., skull and pectoral region. **a**, Left lateral view; **b**, dorsal view with enlargement of scales; and **c**, ventral view with enlargement of anterior ribs. See Fig. 3 for labelled drawing of skull in dorsal view. Abbreviations: an, anocleithrum; bb,

basibranchial; co, coracoid; clav, clavicle; clth, cleithrum; cbr, ceratobranchial; ent, entopterygoid; hu, humerus; lep, lepidotrichia; mand, mandible; nar, naris; or, orbit; psp, parasphenoid; ra, radius; suc, supracleithrum; ul, ulna; uln, ulnare. Scale bar equals 5 cm.

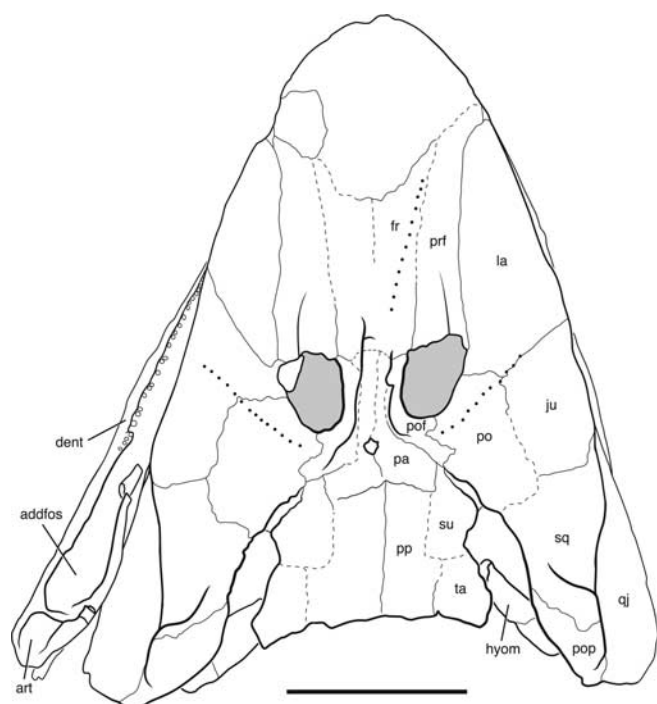


Figure 3 | Drawing of the skull of NUFV 108, holotype of *Tiktaalik roseae* gen. et sp. nov., dorsal view. Dashed lines are unclear contacts, dots are idealized sensory pore lines. Note that the narrow zone between the postorbital and the supratemporal is a slightly separated scarf joint, not a rostral extension of the spiracle. Abbreviations: addfos, adductor fossa; art, articular; dent, dentary; fr, frontal; hyom, hyomandibula; ju, jugal; la, lacrimal; pa, parietal; po, postorbital; pof, postfrontal; pop, preopercular; pp, postparietal; prf, prefrontal; qj, quadratojugal; sq, squamosal; su, supratemporal; ta, tabular. Scale bar equals 5 cm.

indicating that the bodies were briefly exposed before burial. The fossiliferous zone is overlain by a 60-cm-thick featureless siltstone that grades into a palaeosol with root traces and carbonate nodules. The fauna from the site includes the antiarch placoderm *Asterolepis* sp., lungfish, holoptychiid porolepiforms (including *Laccognathus* sp.), osteolepidid and tristichopterid sarcopterygians, and the new elpistostegalian sarcopterygian.

Systematic palaeontology

Sarcopterygii¹⁴
Tetrapodomorpha¹⁵
Elpistostegalia¹⁶

Remarks. Elpistostegalia (=Panderichthyida¹⁷) has generally been used to unite *Elpistostegia* and *Panderichthys* to the exclusion of other sarcopterygians^{5,18}. Most of the features used to support this grouping, however, are also seen in early tetrapods such as *Acanthostega*^{19–22}, *Ichthyostega*²³ and *Ventastega*^{24,25}. Accordingly, a flattened skull with dorsally placed eyes, an elongate snout including paired frontal bones, enlarged prefrontal bone, marginal nares, enlarged spiracle, dorsoventrally flattened body, and loss of the anterior dorsal fin are all attributes of the tetrapod stem lineage and do not unite *Elpistostegia* and *Panderichthys* unambiguously^{26,27}. Consequently, we use Elpistostegalia as the name of the node along the tetrapod stem lineage that includes the common ancestor of *Panderichthys*, *Elpistostegia* and tetrapods and is diagnosed by the characters above. We use the term ‘elpistostegalian fish’ for the paraphyletic grade of flat-headed, finned sarcopterygians that lie along the tetrapod stem lineage.

Tiktaalik roseae gen. et sp. nov.

Locality. Canada, Nunavut Territory, southern Ellesmere Island, near the eastern arm of Bird Fiord, Nunavut Palaeontological Expedition site NV2K17; N77° 09.898' W86° 16.151'.

Horizon. Okse Bay Group, middle part of the Fram Formation.

Age. Late Devonian, early Frasnian stage.

Etymology. *Tiktaalik* (tic'täl'ik) is derived from Inuktitut, the traditional language in Nunavut, and is the name used for a large, freshwater fish seen in the shallows. The species name honours a benefactor of Devonian palaeontology.

Holotype. Nunavut Fossil Vertebrate Collection (NUFV) 108, skull and postcrania (Figs 2 and 3).

Material. This description is based on a suite of specimens (NUFV 108–135) from a single locality (NV2K17). Three specimens (NUFV 108–110) preserve skulls, pectoral girdles and fins in articulation. Consistent features of all elements throughout a twofold size range indicate that this suite of specimens represents a single species. Specimens were mechanically prepared under a binocular microscope. Material is to be housed at the Canadian Museum of Nature, Ottawa, Ontario, until such time as museum and research facilities are developed within the Nunavut Territory.

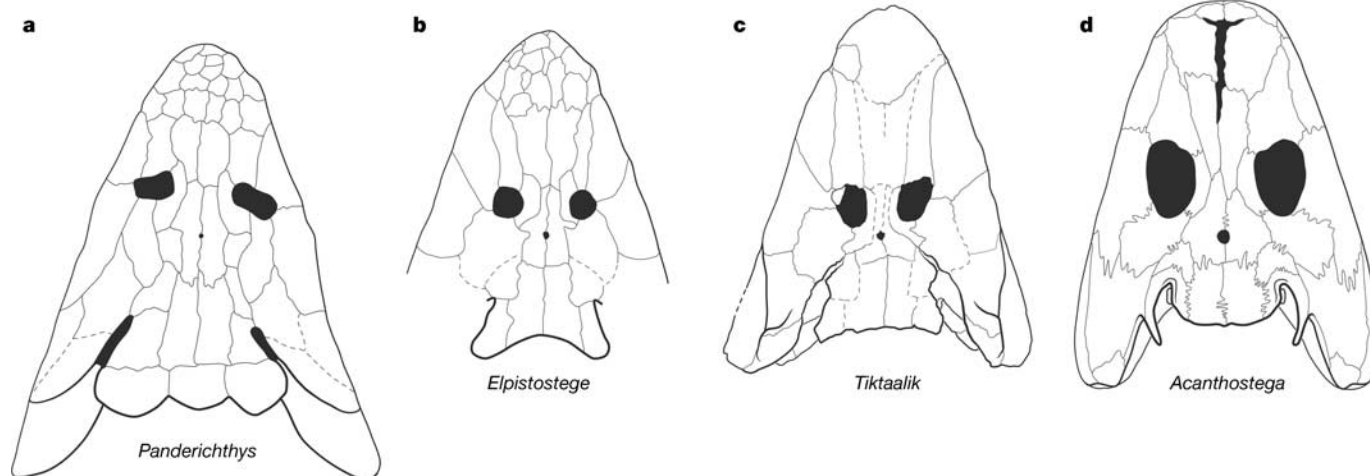


Figure 4 | Skull roofs of elpistostegalian fish and the early tetrapod *Acanthostega*. Skulls drawn to the same snout–postparietal length and aligned at the parietal foramen. Drawings based on published illustrations: *Panderichthys*²⁷, *Elpistostegia*³ and *Acanthostega*²¹. Proportion of snout

(measured as the length anterior to the middle of the orbits relative to the length from the tip of the snout to the posterior margin of the postparietals): a, 38%; b, 58%; c, 62%; d, 55%.

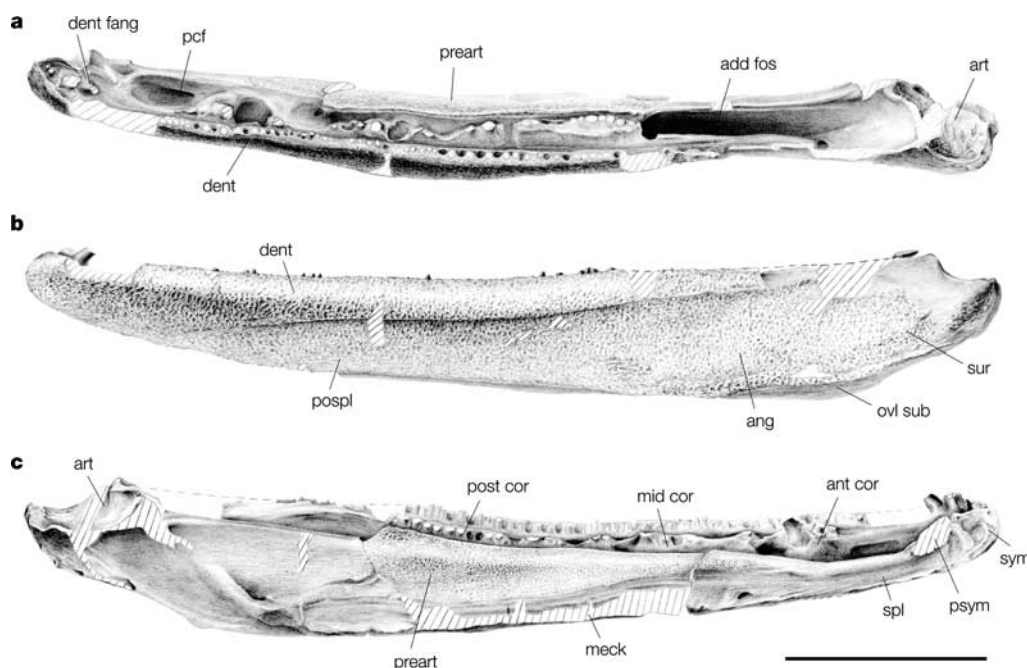


Figure 5 | NUFV 116, left lower jaw of *Tiktaalik roseae*. **a**, Dorsal view; **b**, lateral view; and **c**, medial view. Abbreviations: add fos, adductor fossa; ang, angular; ant cor, anterior coronoid; art, articular; dent, dentary; dent fang, dentary fang; meck, Meckelian bone; mid cor, middle coronoid; ovl

sub, submandibular overlap area; pcf, precoronoid fossa; pospl, postsplenial; post cor, posterior coronoid; preart, prearticular; psym, parasymphysial plate; spl, splenial; sur, surangular; sym, symphysis. Scale bar equals 5 cm.

Diagnosis. Elpistostegalian sarcopterygian differentiated from *Panderichthys* in the loss of the intertemporal, opercular, subopercular and extrascapular bones and in the possession of a snout longer than the postorbital skull roof, a relatively wider spiracular notch, and imbricate ribs. *Tiktaalik* is differentiated from both *Panderichthys* and *Elpistostege* by a postfrontal–supratemporal contact (excluding postorbital–parietal contact), a postfrontal that does not extend anterior to the orbits, and the incorporation of the supratemporal into the medial margin of the spiracular notch. *Tiktaalik* is further differentiated from *Elpistostege* in having narrowly overlapping dorsal scales that are only slightly taller than wide. Among the features that differentiate *Tiktaalik* from *Acanthostega* and other tetrapods are the presence of lepidotrichia in the pectoral and pelvic fins, a relatively elongate hyomandibula, pectoral fin radials that branch, dermal supracleithral elements, a precoronoid fossa in the lower jaw, and a palate with entopterygoids that do not meet at the midline.

Description

The skull of *Tiktaalik* has marginal nares, large prefrontals, and closely spaced, dorsally placed orbits as in *Panderichthys*⁵, *Elpistostege*³ and early tetrapods^{20,23}. The elongation of the rostrum relative to the postorbital region of the skull is more similar to *Elpistostege* and *Acanthostega* than *Panderichthys* (Fig. 4). *Tiktaalik* possesses a broadly overlapping scarf joint between the lacrimal and prefrontal. As in *Panderichthys* and *Elpistostege*, the extent of the lacrimal and jugal along the orbital margin varies, even in the same individual (Fig. 4). The postfrontal is elevated along the medial orbital margin to form a prominent brow ridge, as in *Panderichthys* and *Elpistostege*. A distinct crest in the posterior portion of the cheek is formed by the ventral margins of the squamosal and a small preopercular. Ventral to this crest, an elongate quadratojugal contacts the jugal, a derived condition also seen in *Panderichthys* and early tetrapods. In contrast to *Panderichthys*, the intertemporal is absent in *Tiktaalik*, *Elpistostege* and some early tetrapods. Uniquely, the postfrontal reaches the supratemporal, thereby excluding contact between the postorbital and parietal. In *Panderichthys*, the parietal–postparietal suture is interdigitated⁵, whereas in *Tiktaalik* the suture is linear and slightly

separated in NUFV 108 and NUFV 110. There is a firm sutural union between supratemporal and parietal.

The spiracular notch in *Tiktaalik* is wider than that in *Panderichthys* and comparable in width to that in Devonian tetrapods such as *Ventastega*²⁵. The width of the spiracular notch is unknown in *Elpistostege*. In *Tiktaalik*, the notch is bounded medially by the tabular and supratemporal; *Panderichthys*, in contrast, does not incorporate the supratemporal into the medial margin of the spiracular notch, and a reconstruction of *Elpistostege*²⁷ appears to show the same condition. In *Tiktaalik*, the supratemporal–postorbital contact is a scarf joint extending anteriorly from the spiracular notch. The morphology of the sutural contact between these elements in *Panderichthys* and *Elpistostege* is unclear. The rugose texture of the postparietal shield of *Tiktaalik* is suggestive of muscle insertion over much of this surface except for ornamented portions of the tabular and raised margins of the sutures.

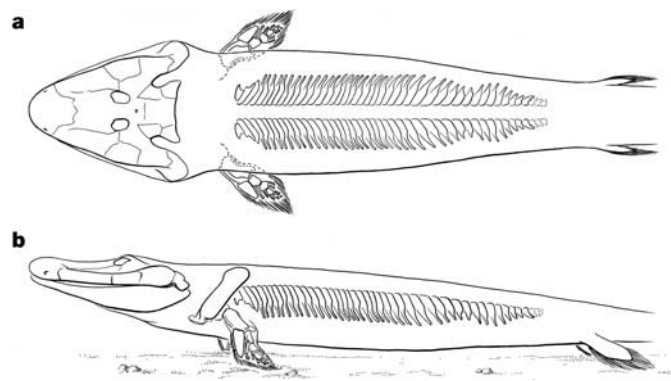


Figure 6 | Interpretive skeletal reconstruction of *Tiktaalik*. **a**, Dorsal view of body with scales removed depicting orientation of ribs as preserved in NUFV 108. **b**, Lateral view. The ribs are shown in dorsoventral orientation. Number of ribs is estimated from the incompletely preserved series in NUFV 108.

In NUFV 108, the exposed portion of the palate reveals that the entopterygoids are separated by the parasphenoid and vomers anteriorly and do not meet in the midline (Fig. 2c), as in *Panderichthys*⁵ and in contrast to the inter-entopterygoid contact of early tetrapods. The posterior part of the palatoquadrate in NUFV 108 is a dorsoventrally shallow ramus, as in *Panderichthys* and early tetrapods¹⁰. The remainder of the palate in specimens of *Tiktaalik* is not known because it is obscured by lower jaws, gulars and branchial elements.

The sample of *Tiktaalik* lower jaws, with lengths ranging from 170–310 mm, represents at least ten individuals. The configuration of the bony elements comprising the lower jaw and the arrangement of the dentition are similar to that in *Panderichthys*²⁸ (Fig. 5). The rostral end of the lower jaw of *Tiktaalik* is not as deep as that of *Panderichthys*. There is no indication of dorsal closure of the precoronoid fossa nor the starburst pattern of ornament on the infradentaries that characterize early tetrapod jaws²⁸.

The elongate and robust ceratobranchials in *Tiktaalik* extend into the gill chamber and bear a deep, longitudinal vascular sulcus along their ventral surfaces that is indicative of well-developed gills²⁹. The short, piriform hyomandibula spans the spiracular notch, articulates medially with facets on the braincase, and attenuates laterally at its contact with the medial surface of the palatoquadrate.

The three articulated specimens of *Tiktaalik* (NUFV 108–110) that preserve skulls, pectoral girdles and fins show no indication of an opercular or subopercular, despite the relatively complete and articulated preservation of other skeletal elements in the shoulder and posterior skull region. In NUFV 108 (Fig. 2), which was preserved dorsal-side down, the complete and largely undistorted preservation of the dorsal skull roof, pectoral girdles and dorsal scales adjacent to the area where opercular and subopercular would lie supports the inference that the absence of these bones is not due to postmortem disassociation. The extrascapular series is also absent, eliminating bony connection between the skull and the pectoral girdle. Submandibulars and large gulars are present in *Tiktaalik*; the status of the preoperculosubmandibular is indeterminate.

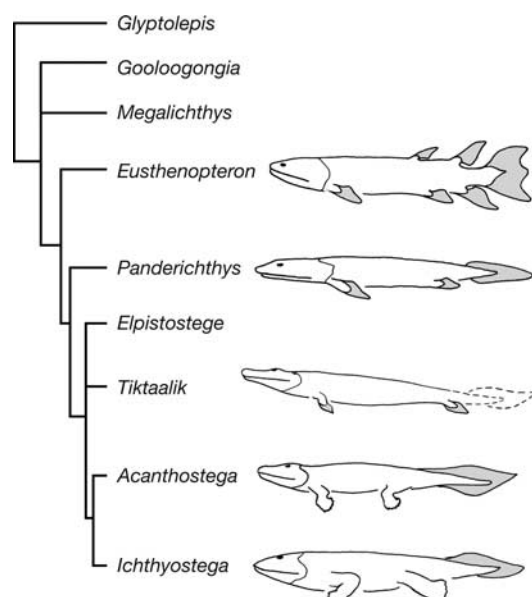


Figure 7 | Strict consensus tree from a phylogenetic analysis of 114 characters and nine taxa. *Tiktaalik* is the sister group of *Acanthostega*+*Ichthyostega* in one of the two most parsimonious trees, and clades with *Elpistostege* as sister to the tetrapods in the other. Tree length = 149, consistency index = 0.8389, consistency index excluding uninformative characters = 0.7966, retention index = 0.8140, and rescaled consistency index = 0.6828. The characters list and data matrix are available as Supplementary Information.

The dorsal lamina of the cleithrum is posteriorly inclined and thus presents a low profile in lateral view as in *Panderichthys*. The robust cranial margin of the cleithrum forms a buttress for the posterior margin of the gill chamber. The ornamented supracleithral series is complete; a posttemporal overlaps the supracleithrum, which in turn overlaps an elongate anocleithrum. The presence of an ornamented supracleithral series is shared with more primitive sarcopterygian fish whereas the loss of the extrascapular series was previously known only in tetrapods. Details of the shoulder girdle and fins are described in an accompanying paper³⁰.

The dorsal surface of *Tiktaalik* is covered with rhombic, overlapping, tuberculated scales that are similar to those of *Panderichthys* (Fig. 2b). In both taxa, the scales are slightly taller than wide and each scale narrowly overlaps the adjacent posteroventral scale. In contrast, the rhombic scales of *Elpistostege* are twice as tall as wide and overlap the adjacent scale over a much broader area². As in *Panderichthys*, the dorsal scale cover in *Tiktaalik* (NUFV 108) shows no interruption for an anterior dorsal fin.

The vertebral elements appear to be unossified in NUFV 108 (Fig. 2). In contrast, the well-ossified costal skeleton indicates a presacral vertebral count of approximately 45 (Fig. 6), which is higher than *Eusthenopteron*³¹ and *Acanthostega*²² (about 30), or *Ichthyostega*³² (about 26). The ribs of *Tiktaalik* (Fig. 6) bear plate-like flanges that extend from the caudal margin of each shaft. Distally the ribs attenuate to conventional rod-like form. Posterior ribs of the trunk are shorter, and the flanges are broadly triangular. Adjacent ribs imbricate with the internal surfaces overlapping the external surface of the rib behind. The anterior-most ribs exposed in NUFV 108 have ventrally reflected uncinate processes along their cranial margins (Fig. 2c). The expanded, imbricate ribs of *Tiktaalik* extend more ventrally and are broader than the ribs reported for *Panderichthys*⁵, which do not imbricate. Imbricate ribs were previously known only in tetrapods such as *Ichthyostega*^{23,32}.

Phylogenetic relationships

A phylogenetic analysis of sarcopterygian fishes and early tetrapods (Fig. 7) supports the hypothesis that *Tiktaalik* is the sister group of tetrapods or shares this position with *Elpistostege*. *Tiktaalik* retains primitive tetrapodomorph features such as dorsal scale cover, paired fins with lepidotrichia, a generalized lower jaw, and separated entopterygoids in the palate, but also possesses a number of derived features of the skull, pectoral girdle and fin, and ribs that are shared with stem tetrapods such as *Acanthostega* and *Ichthyostega*. *Tiktaalik* is similar to these forms in the possession of a wide spiracular tract and



Figure 8 | Palaeogeographic reconstruction of Euramerican landmass during the Late Devonian. We redraw this figure from 'Palaeogeographic globes—Late Devonian (2001)' with the permission of R. Blakey (http://jan.ucc.nau.edu/~rcb7/Late_Dev.jpg). Palaeoequator marked by dashed line. Sites of elpistostegalian fish discoveries: T, *Tiktaalik*; E, *Elpistostege*; P, *Panderichthys*.

the loss of the opercular, subopercular and extrascapulars. The pectoral girdle is derived in the degree to which the scapulocoracoid is expanded dorsally and ventrally, and the extent to which the glenoid fossa is oriented laterally³⁰. The pectoral fin is apomorphic in the elaboration of the distal endoskeleton, the mobility of segmented regions of the fin, and the reduction of lepidotrichia distally³⁰.

Panderichthys, *Elpistostege* and *Tiktaalik* are a paraphyletic assemblage of elpistostegalian fish along the tetrapod stem that lack the anterior dorsal fins and possess broad, dorsoventrally compressed skulls with dorsally placed eyes, paired frontal bones, marginal nares, and a subterminal mouth. The paraphyletic nature of this assemblage is particularly informative for reconstructing features of the common ancestor of tetrapods and unravelling the sequence of character acquisition in the evolution of the tetrapod body plan. Derived features in the rear of the skull preceded changes to the lower jaws and palate. In addition, loss of the opercular, subopercular and the extrascapular series occurred before the loss of the gulars, submandibulars and the supracleithral series. Some tetrapod-like features evolved independently in other sarcopterygian groups. For example, apomorphies of the fin and shoulder of *Sauripterus*³³ and the neck region of *Mandageria*³⁴ are homoplasies that are shared with more basal members of the tetrapod stem.

Biogeography of *Tiktaalik*

The temporal and geographic provenance of *Panderichthys*, *Elpistostege* and *Tiktaalik* indicate that the radiation of elpistostegalian fish occurred during the late Givetian to early Frasnian within the Euramerican landmass (Fig. 8), strengthening the argument for the origin of tetrapods within Euramerica³⁵. The fauna associated with *Tiktaalik* is very similar to that from the late Givetian to early Frasnian deposits at the Lode quarry in Latvia that produce *Panderichthys*, thus providing strong evidence of a palaeobiogeographic connection between Nunavut and the Baltic region during the Late Devonian. The non-marine depositional setting for *Tiktaalik* differs from the deltaic and estuarine settings of *Panderichthys*³⁶ and *Elpistostege*³⁷, respectively, suggesting that elpistostegalian fishes were exploiting a range of habitats.

Palaeobiology of *Tiktaalik*

Overall, the skeleton of *Tiktaalik* is that of a flat-bodied animal with raised and dorsally placed eyes, a mobile neck, imbricate ribs, and a pectoral girdle and forefin capable of complex movements and substrate support³⁰ (Fig. 6). This suite of features represents a major departure from the pattern in more primitive sarcopterygian fishes. The sedimentological interpretation of the Fram Formation indicates that *Tiktaalik* lived in a low gradient, meandering fluvial system within a subtropical to tropical climatic belt¹³. In this setting, *Tiktaalik* developed new mechanisms of head movement, respiration and body support that enabled this fish to exploit shallow water and even subaerial habitats. In support of this interpretation, ribs of the type that occur in *Tiktaalik* augment thoracolumbar rigidity and axial support³⁸, functions that are not necessary in an aquatic setting that is deep enough to support the body.

Many of the specialized features of *Tiktaalik* relate to changes in respiration relative to that in more primitive sarcopterygians. *Tiktaalik* is transitional in the evolutionary shift from the pharyngeal and opercular pumps employed by fish to the buccal and costal pumping mechanisms of tetrapods. The expanded gular plates and robust branchial elements could have provided the mechanical basis for buccal pumping for lungs as well as gills. The emphasis on buccal pumping is further augmented by the expansion of the width of the skull, which enhances the volume of the buccal cavity. The likelihood that gular plates and other branchial elements assumed a predominant respiratory function for air breathing in *Tiktaalik* is increased by loss of the operculum and the apparent reduction of the opercular apparatus.

An enlarged spiracle has been interpreted to have a respiratory role

in early tetrapods^{10,35} and modifications to the spiracular region across the fish–tetrapod transition may be specializations for spiracular breathing across the aquatic–terrestrial interface. The condition of the hyomandibula and spiracular tract in *Tiktaalik* illustrates part of the phylogenetic trajectory of enlarging the spiracular tract and reducing the hyomandibula. In comparison to osteolepiforms, *Panderichthys* has a relatively wide spiracle and a shortened, rod-like hyomandibula ending at the opercular process¹⁰. *Tiktaalik* continues this trend with a still wider, tetrapod-like spiracular tract and a shorter, more robust hyomandibula. In *Acanthostega*, the broad spiracular tract houses a short, robust stapes with the stapedial footplate lodged in the fenestra vestibuli of the braincase³⁹, although the stapes retains its primitive role in palatal and spiracular movements⁴⁰.

These changes in cranial architecture are also associated with new patterns of locomotion³⁰ and, apparently, feeding. The loss of the opercular, subopercular and the extrascapular series effectively decoupled the head from the pectoral girdle, introducing an independent range of motion of the head, and greater freedom of the pectoral girdle and fin. An extensive area for cervical muscle insertion is developed across the dorsal surface of the posterior skull roof, augmenting the potential for head mobility. The robust scapulocoracoid, extensive endochondral skeleton and reduced lepidotrichia of the pectoral fin suggest a strong and flexible appendage capable of complex movements across a substrate³⁰.

Major elements of the tetrapod body plan originated as a succession of intermediate morphologies that evolved mosaically and in parallel among sarcopterygians closely related to tetrapods, allowing them to exploit diverse habitats in the Devonian. The geological setting in which *Tiktaalik* was found supports the view that shallow water habitats on Late Devonian floodplains of the Euramerican landmass were the locus for the fish–tetrapod transition. New discoveries of transitional fossils such as *Tiktaalik* make the distinction between fish and the earliest tetrapods increasingly difficult to draw.

Received 11 October 2005; accepted 8 February 2006.

- Westoll, T. S. Ancestry of the tetrapods. *Nature* **141**, 127 (1938).
- Schultze, H.-P. & Arsenault, M. The panderichthyid fish *Elpistostege*: A close relative of tetrapods? *Palaeontology* **28**, 292–309 (1985).
- Schultze, H.-P. in *Devonian Fishes and Plants of Miguasha, Quebec, Canada* (eds Schultze, H.-P. & Cloutier, R.) 316–327 (Friedrich Pfeil, München, 1996).
- Vorobyeva, E. I. Observations on two rhipidistian fishes from the Upper Devonian of Lode, Latvia. *Zool. J. Linn. Soc.* **70**, 191–201 (1980).
- Vorobyeva, E. I. & Schultze, H.-P. Description and systematics of panderichthyid fishes with comments on their relationship to tetrapods. In *Origins of the Higher Groups of Tetrapods* (eds Schultze, H.-P. & Trueb, L.) 68–109 (Cornell Univ. Press, Ithaca, 1991).
- Vorobyeva, E. I. The shoulder girdle of *Panderichthys rhombolepis* (Gross) (Crossopterygii), Upper Devonian, Latvia. *Geobios* **19**, 285–288 (1995).
- Vorobyeva, E. I. Morphology of the humerus in the Rhipidistian Crossopterygii and the origin of tetrapods. *Paleontol. J.* **34**, 632–641 (2000).
- Ahlberg, P. E., Clack, J. A. & Luksevics, E. Rapid braincase evolution between *Panderichthys* and the earliest tetrapods. *Nature* **381**, 61–64 (1996).
- Boisvert, C. A. The pelvic fin and girdle of *Panderichthys* and the origin of tetrapod locomotion. *Nature* **438**, 1145–1148 (2005).
- Brazeau, M. D. & Ahlberg, P. E. Tetrapod-like middle ear architecture in a Devonian fish. *Nature* **439**, 318–321 (2006).
- Embry, A. & Klován, J. E. The Middle-Upper Devonian clastic wedge of the Franklinian Geosyncline. *Bull. Can. Petrol. Geol.* **24**, 485–639 (1976).
- Chi, B. I. & Hills, L. V. Biostratigraphy and taxonomy of Devonian megaspores. *Bull. Can. Petrol. Geol.* **24**, 640–813 (1976).
- Embry, A. F. in *Geology of the Inuitian Orogen and Arctic Platform of Canada and Greenland* (ed. Trettin, H. P.) 263–279 (Geological Survey of Canada, Calgary, 1991).
- Romer, A. S. Herpetichthyes, Amphibioidei, Choanichthyes, or Sarcopterygii? *Nature* **176**, 126 (1955).
- Ahlberg, P. E. A re-examination of sarcopterygian interrelationships, with special reference to the Porolepiformes. *Zool. J. Linn. Soc.* **103**, 241–287 (1991).
- Camp, C. L. & Alison, H. J. Bibliography of fossil vertebrates 1949–1953. *Geol. Soc. Am. Mem.* **84**, 1–53 (1961).
- Vorobyeva, E. I. Panderichthyida—A new order of Paleozoic crossopterygian fishes (Rhipidistia). *Dokl. Akad. Nauk SSR* **306**, 188–189 (1989).

18. Janvier, P. *Early Vertebrates* 1–393 (Clarendon, Oxford, 1996).
19. Clack, J. A. *Acanthostega gunnari*, a Devonian tetrapod from Greenland: the snout, palate and ventral parts of the braincase, with a discussion of their significance. *Medd. Groen. Geosci.* **31**, 1–24 (1994).
20. Clack, J. A. The dermal skull roof of *Acanthostega gunnari*, an early tetrapod from the Late Devonian. *Trans. R. Soc. Edinb. Earth Sci.* **93**, 17–33 (2002).
21. Clack, J. A. A revised reconstruction of the dermal skull roof on *Acanthostega gunnari*, an early tetrapod from the Late Devonian. *Trans. R. Soc. Edinb. Earth Sci.* **93**, 163–165 (2003).
22. Coates, M. I. The Devonian tetrapod *Acanthostega gunnari* Jarvik: Postcranial anatomy, basal tetrapod interrelationships and patterns of skeletal evolution. *Trans. R. Soc. Edinb. Earth Sci.* **87**, 363–421 (1996).
23. Jarvik, E. The Devonian tetrapod *Ichthyostega*. *Fossils Strata* **40**, 1–213 (1996).
24. Ahlberg, P. E., Luksevics, E. & Lebedev, O. The first tetrapod finds from the Devonian (Upper Famennian) of Latvia. *Phil. Trans. R. Soc. Lond. B* **343**, 303–328 (1994).
25. Luksevics, E. & Zupins, I. Sedimentology, fauna, and taphonomy of the Pavari site, Late Devonian of Latvia. *Acta Univ. Latviensis* **679**, 99–119 (2004).
26. Cloutier, R. & Ahlberg, P. E. in *Interrelationships of Fishes* (eds Stiassny, M. L. J., Parenti, L. R. & Johnson, G. D.) 445–479 (Academic, New York, 1996).
27. Ahlberg, P. E., Luksevics, E. & Mark-Kurik, E. A near tetrapod from the Baltic Middle Devonian. *Palaeontology* **43**, 533–548 (2000).
28. Ahlberg, P. E. & Clack, J. Lower jaws, lower tetrapods—A review based on the Devonian genus *Acanthostega*. *Trans. R. Soc. Edinb. Earth Sci.* **89**, 11–46 (1998).
29. Coates, M. I. & Clack, J. Fish-like gills and breathing in the earliest known tetrapod. *Nature* **352**, 234–236 (1991).
30. Shubin, N. H., Daeschler, E. B. & Jenkins, F. A. Jr. The pectoral fin of *Tiktaalik roseae* and the origin of the tetrapod limb. *Nature* doi:10.1038/nature04637 (this issue).
31. Andrews, S. M. & Westoll, T. S. The postcranial skeleton of *Eusthenopteron foordi* Whiteaves. *Trans. R. Soc. Edinb.* **68**, 207–329 (1970).
32. Ahlberg, P. E., Clack, J. A. & Blom, H. The axial skeleton of the Devonian tetrapod *Ichthyostega*. *Nature* **437**, 137–140 (2005).
33. Davis, M. C., Shubin, N. H. & Daeschler, E. B. A new specimen of *Sauripterus taylori* (Sarcopterygii, Osteichthyes) from the Famennian Catskill Formation of North America. *J. Vert. Paleontol.* **24**, 26–40 (2004).
34. Johanson, Z., Ahlberg, P. & Ritchie, A. The braincase and palate of the tetrapodomorph sarcopterygian *Mandageria fairfaxi*: Morphological variability near the fish–tetrapod transition. *Palaeontology* **46**, 271–293 (2003).
35. Clack, J. A. *Gaining Ground: The Origin and Evolution of Tetrapods* 1–369 (Indiana Univ. Press, Bloomington, 2002).
36. Upeniece, I. The unique fossil assemblage from the Lode Quarry (Upper Devonian, Latvia). *Mitt. Mus. Natkd. Berl. Geowiss.* **4**, 101–119 (2001).
37. Chidiac, Y. in *Devonian Fishes and Plants of Miguasha, Quebec, Canada* (eds Schultze, H.-P. & Cloutier, R.) 47–53 (Friedrich Pfeil, Munchen, 1996).
38. Jenkins, F. A. Jr. Anatomy and function of expanded ribs in certain edentates and primates. *J. Mammal.* **51**, 288–301 (1970).
39. Clack, J. A. Earliest known tetrapod braincase and the evolution of the stapes and fenestra ovalis. *Nature* **369**, 392–394 (1994).
40. Clack, J. A. Discovery of the earliest known tetrapod stapes. *Nature* **342**, 425–427 (1989).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The illustrations are the work of K. Monoyios (Figs 1, 5, 6, 7) and K. Luckenbill (Figs 2, 3, 4, 8). Specimens were prepared by C. F. Mullison. J. P. Downs discovered the NV2K17 locality. S. Gatesy discovered and excavated NUFV 108. This research was supported by permits from the Nunavut Ministry of Culture, Languages, Elders and Youth (D. Stenton, J. Ross), the Grise Fiord Hamlet and HTA, with logistical support from the Polar Continental Shelf Project (B. Hycryk, D. Maloley, J. MacEachern, W. Benoit, G. Benoit, N. Couture, H. Gordon, D. Clouthier, D. Mueller), and collections support from the Canadian Museum of Nature (S. Cumbaa, K. Shepherd). The elders of the Inuit Qaujimajatuqangit Katimajit of the Government of Nunavut provided the name for the new genus. A. Embry and U. Mayr provided input at the conception of the field project. M. Friedman and M. Coates supplied insights. Assistance in the field (1999–2004) was provided by W. Amaral, B. Atagootak, J. Conrad, M. Davis, J. Downs, S. Gatesy, S. Madsen, K. Middleton, K. Monoyios, C. Schaff, M. Shapiro, R. Shearman and C. Sullivan. This research was supported by a patron of our research, the Academy of Natural Sciences, the Putnam Expeditionary Fund (Harvard University), the University of Chicago, the National Science Foundation, and the National Geographic Society Committee for Research and Exploration.

Author Contributions E.B.D. and N.H.S. conceived and co-directed the project. F.A.J. Jr collaborated on all phases of the research.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to E.B.D. (daeschler@acnatsci.org) or N.H.S. (nshubin@uchicago.edu).

ARTICLES

The pectoral fin of *Tiktaalik roseae* and the origin of the tetrapod limb

Neil H. Shubin¹, Edward B. Daeschler² & Farish A. Jenkins Jr³

Wrists, ankles and digits distinguish tetrapod limbs from fins, but direct evidence on the origin of these features has been unavailable. Here we describe the pectoral appendage of a member of the sister group of tetrapods, *Tiktaalik roseae*, which is morphologically and functionally transitional between a fin and a limb. The expanded array of distal endochondral bones and synovial joints in the fin of *Tiktaalik* is similar to the distal limb pattern of basal tetrapods. The fin of *Tiktaalik* was capable of a range of postures, including a limb-like substrate-supported stance in which the shoulder and elbow were flexed and the distal skeleton extended. The origin of limbs probably involved the elaboration and proliferation of features already present in the fins of fish such as *Tiktaalik*.

A landmark event in vertebrate history is the transformation of fish fins into tetrapod limbs. Insights into this transition illuminate the biological mechanisms that generate major shifts in developmental genetics^{1–6}, skeletal structure^{7,8} and biomechanics^{9–11}. Limb skeletons differ from those of fins mainly by the presence of bones that comprise mobile wrists, ankles and digits. In addition, limbs lack the extensive array of lepidotrichia, which are dermal rods that form much of the surface area of fins.

An impediment to understanding the fin–limb transition has been the nature of available evidence from the sister group of tetrapods. The closest living relatives of tetrapods—lungfishes and coelacanths—either lack homologous elements to distal limb bones or are so specialized that comparisons with tetrapods are uncertain. Sarcopterygian taxa with possibly intermediate morphologies are restricted to fossil forms from the Palaeozoic era. A Devonian rhizodontid, *Sauripterus*, is known to possess digit-like radials, but phylogenetic analyses indicate that this group is not the closest relative of tetrapods^{12–14}. The current hypothesis is that the sister group of tetrapods are elpistostegids^{15,16}, sarcopterygians known from Quebec and Latvia^{17–20}. Unfortunately, the distal region of the best-known pectoral fin of the elpistostegid *Panderichthys* is covered by lepidotrichia and the complete distal endoskeleton is unknown¹¹. A strict interpretation of this taxon has led to proposals that the distinctive features of tetrapod limbs are novelties, with few antecedents in fish fins^{1–6,8}. If this scenario is true, then the origin of tetrapods involved major changes in skeletal patterning and appendage function. Palaeontological data thus have a uniquely important role in interpreting the origin of limbs.

The discovery of *Tiktaalik roseae*²¹ brings new data to bear on these issues. The material is remarkable for its phylogenetic position, three dimensional preservation and abundance. As a member of the sister group of limbed vertebrates, *Tiktaalik* can reveal the primitive pattern from which limbs were derived. Articulated pectoral fins have been prepared from three different specimens, two of which preserve the anatomical relations of endochondral bones, lepidotrichia and scales (Figs 1 and 2; see also Supplementary Information). There is little variation in articular design or appendage proportions despite a twofold difference in size between the smallest (Nunavut

Fossil Vertebrate Collection (NUFV) 108) and largest (NUFV 109) fins. Left and right fins of NUFV 109 were disarticulated and individual bones prepared free to expose the joint surfaces and articular relationships (Supplementary Information). Multiple specimens record details of bony anatomy and articular geometry for major elements of the appendage, providing a high degree of confidence in the fin reconstruction and the interpretation of joint function (Fig. 2).

In overall structure, the pectoral girdle and fin of *Tiktaalik* are distinguished from those of other tetrapodomorph fish in having an expanded endoskeleton and a relatively reduced dermal exoskeleton (Figs 1–3; see also Supplementary Information). The lepidotrichia are solid rods that remain unjointed for most of their length and invest the endochondral bones dorsally and ventrally (Fig. 1; see also Supplementary Information). The lepidotrichial sheath is heterogeneous; lepidotrichia that cover anterior endochondral bones are more robust and originate more proximally than those on the posterior side (Fig. 1a). Along the fringe of the fin distal to the last endochondral radials these rods are extensively jointed. The fin is relatively stouter and anteroposteriorly narrower than the fins of other tetrapodomorph fishes (Fig. 4).

Comparative anatomy of the pectoral girdle

The shoulder of *Tiktaalik* is highly derived with a suite of features only known in tetrapods and *Panderichthys*. The endochondral components (scapula and coracoid) of the pectoral girdle are enlarged and the dermal series (cleithrum, clavicle, anocleithrum and supracleithrum) is relatively reduced (Figs 3 and 5a, b). Scapular height is increased relative to that of tristichopterids via a broad process that extends dorsal to the glenoid to lie flush against the medial aspect of the cleithrum, as in rhizodontids (Figs 3 and 5b). The glenoid is oriented posteroventrolaterally and partially exposed in lateral view, which is intermediate between the posterior orientation of the glenoid in *Eusthenopteron*⁹ and the lateral orientation of *Acanthostega*²² and other basal tetrapods²³ (Figs 3 and 5a). In addition, the glenoid lies adjacent to the ventral surface of the coracoid, and accordingly the fin projects from the body near the level of the belly, as in *Panderichthys*²⁴ (Figs 3 and 5a, b). The coracoid

¹Department of Organismal Biology and Anatomy, The University of Chicago, Chicago, Illinois 60637, USA. ²Academy of Natural Sciences of Philadelphia, 19th and Benjamin Franklin Parkway, Philadelphia, Pennsylvania 19103, USA. ³Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, USA.

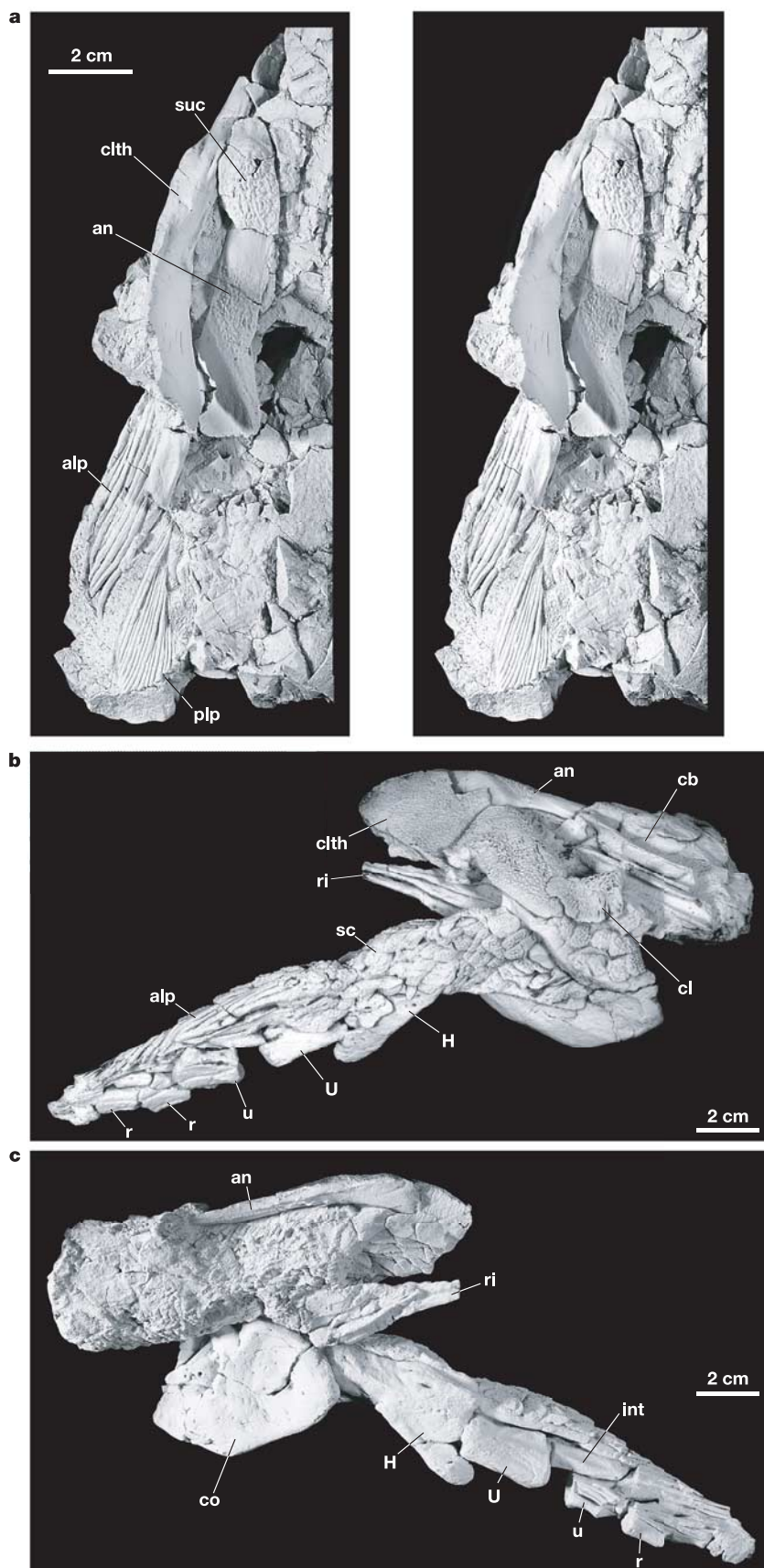


Figure 1 | Articulated pectoral fins of *Tiktaalik*

roseae. **a**, Stereo pair of left pectoral fin of NUFV 108 in dorsal view showing disparity in size and position of anterior (alp) and posterior (plp) unjointed lepidotrichia and the relative position of dermal girdle elements. **b**, Right pectoral fin of NUFV 110 in anterior view showing preservation of anterior lepidotrichia (alp), clavicle (cl), scales (sc) and endochondral bones in articulation (H, humerus; U, ulna; u, ulnare; r, radials). The anterior lepidotrichia terminate at the elbow, thus allowing a full range of flexion at that joint. **c**, Right pectoral fin of NUFV 110 in ventral view showing positions of coracoid (co) and endochondral and dermal fin elements. an, anocleithrum; cb, ceratobranchial; clth, cleithrum; int, intermedium; ri, rib; suc, supracleithrum.

is a ventrally positioned, horizontally oriented plate that is perforated by a large foramen with a diameter of approximately one-quarter the mediolateral width of the bone (Figs 1c and 3a, b; see also Supplementary Information). A coracoid foramen of comparable size is only known in *Panderichthys*²⁴. The prominent external and internal rims of the coracoid foramen are each interrupted by a sulcus (Supplementary Information). The ventromedial sulcus of the internal rim is aligned with the dorsolateral sulcus of the external rim, an arrangement that is compatible with the transmission of a neurovascular or muscular structure. We interpret the sulci as indicative of a musculotendinous bundle that extends from an origin on the dorsomedial surface of the coracoid plate to an insertion on the ventral surface of the humerus. The coracoids of *Tiktaalik* and *Panderichthys* are generally more massive and thicker than those of either *Acanthostega*²² or *Hynerpeton*²³. Whereas the cleithrum, anclethrum and supracleithrum of *Tiktaalik* are reduced relative to those in other tetrapodomorph fish, the ornamentation on their exposed surfaces is a primitive retention. Notably, the dermal series is reduced to the point where the pectoral girdle has no bony connection with the skull—the extrascapular series, suboperculum and operculum are absent²¹.

Comparative anatomy of the pectoral fin

The endochondral bones of the pectoral fin of *Tiktaalik* combine features of *Eusthenopteron* and *Acanthostega*, and in some aspects are

intermediate. A robust ventral ridge, similar to that in *Eusthenopteron*⁹, extends diagonally across the long axis of the humerus, from the anterior margin of the humeral head to the distal tip of the entepicondyle (Figs 2 and 5c). The leading margin of the humerus is a narrow, strap-shaped surface comparable to those in *Acanthostega*²² and ANSP 21350 (ref. 25). On the dorsal surface, the ectepicondyle is a prominent, distally directed process that extends to the level of the ulnar and radial facets (Fig. 2), much as in basal tetrapods but in contrast to *Eusthenopteron*. The radius is elongate, tapered distally and slightly cambered ventrally (Fig. 2). Unlike the radii of tetrapods and other tetrapodomorphs, with the exception of *Panderichthys*¹¹, the leading margin of the radius forms a sharp crest, whereas the posterior edge is sub-cylindrical. The ulna retains a primitive cuboidal shape and lacks the olecranon process of tetrapods. The ulna resembles that in tetrapods and *Sauripterus*¹⁴ in the absence of a postaxial process and in the presence of multiple articular facets for distal radials (Figs 2, 4 and 6d). Unlike *Sauripterus*, these facets are evidence of synovial joints that promote inter-osseous movement. In *Tiktaalik*, the ulna articulates with four proximal radials, whereas the intermedium articulates with only one (Figs 2 and 6d). An important functional aspect of the articulations is that the five radials share a common axis of flexion/extension that runs anteroposteriorly across the fin. The third radial in the proximal series is enlarged and supports three intermediate radials at joints that are also aligned anteroposteriorly. The central radial of this triplet is, itself, enlarged and articulates with two distal radials. Accordingly, the ulna and enlarged proximal and intermediate radials establish a central axis to the fin, a feature common to basal sarcopterygians^{7,26} but unknown in tetrapodomorphs.

Functional anatomy

An assessment of joint excursions and fin postures and functions in *Tiktaalik* is made possible by the preservation, in multiple specimens, of complete bony articular surfaces of all the major joints of the shoulder and fin (Fig. 6). The excursions possible at each joint may be inferred from the articular geometry, in particular the curvatures of apposing cartilage-covered surfaces. Although this soft tissue is not preserved, the curvature is unlikely to have been less than that preserved as the bony supporting surface of the joint, and could possibly have been greater.

Most of the glenohumeral joint is composed of a large hemispherical humeral head articulating with an ovoid, concave glenoid (Fig. 6a). Typically, this kind of joint geometry permits three degrees of freedom: rotation (supination/pronation), flexion/extension (elevation/depression) and protraction/retraction. In the case of *Tiktaalik*, however, some restriction of humeral mobility is engendered by cranial extensions of the articular surfaces of the glenoid and humeral head (Figs 3d and 6a; see also Supplementary Information). These accessory facets—a shallowly concave, elongate humeral facet and a convex glenoid facet—are brought into contact as the humerus is pronated, flexed and protracted. Pronation, flexion and protraction are movements that could have been effected by the large musculotendinous apparatus passing posterolaterally through the coracoid foramen and inserting on the ventral surface of the humerus. The simultaneous apposition of the reversed concavoconvex geometries of the anterior and posterior parts of the articulation represents a close-packed, or most stable, joint position. Additional stability would be contributed through the action of the trans-coracoid musculature. With the anterior facets of the glenohumeral joint in full contact, however, protraction and supination are inhibited, and flexor forces simply compress and stabilize the humerus against the anterior portion of the glenohumeral joint.

The joints of the elbow provide evidence of the independent mobility of the radius and ulna (Figs 6b and 7). As in *Acanthostega*²², the humeral facets for the radius and ulna are separate from one another, in contrast to the confluent facets in *Eusthenopteron*⁹ and

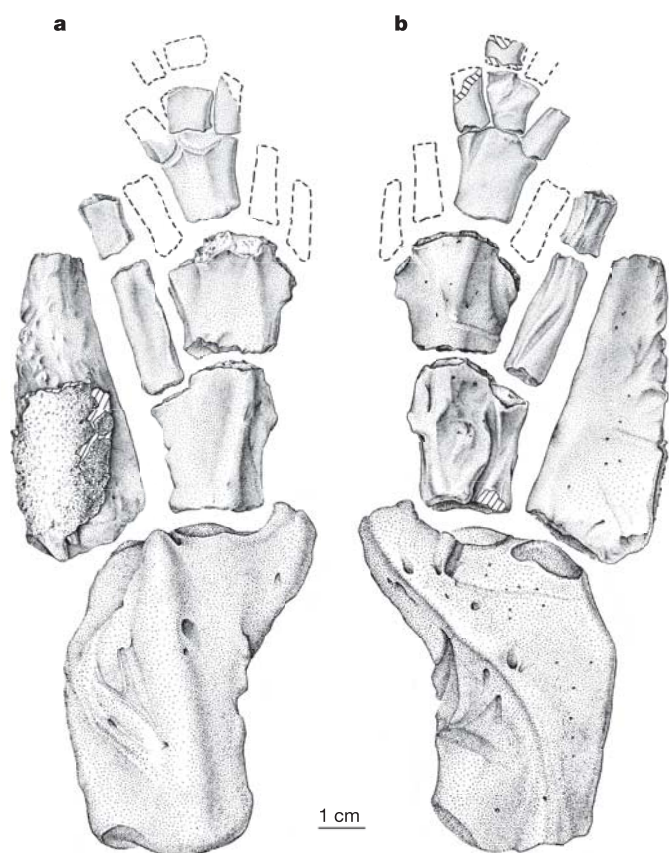


Figure 2 | Reconstruction of the right pectoral fin of *Tiktaalik*. **a**, Dorsal view; **b**, ventral view. Elements with stipple shading were preserved in articulation in NUFV 109 and prepared in the round. Elements with a dashed outline are reconstructed based on their presence in the articulated distal fin of NUFV 110. It is not known how many radials lie distal to the first, second and fourth in the proximal series. Note the dorsal expansion of the distal articular facets on the ulna and third distal radial/mesomere. The dorsal expansion of these facets would have facilitated extension of the distal fin.

other non-tetrapodomorph fishes²⁵. The radial facet is smaller than and anteroventral to the ulnar facet. Neither facet faces exactly distally; both are slightly ventrally oriented, with the radial facet more so than the ulnar. On this evidence the antebrachium would have been slightly flexed in the rest position, with the radius more flexed than the ulna.

Joint geometry provides evidence of the movements possible at the elbow. In outline the convex radial facet of the humerus is a bent ellipsoid; the apposing facet on the radius is a simple, oval concavity (Fig. 6b). Given these shapes, the radius was not capable of independent rotation on the humerus without conjunct translation. The radial facet forms a curved pathway on the humerus: the anterior portion of the facet's surface is nearly in the plane of the ventral surface of the humerus, whereas the posterior part of the facet faces more distally. Accordingly, translation of the radius along this facet provides for a small degree of pronation and supination. The bulbous ulnar facet of the humerus is principally oval in outline and hemispheroidal. The humeral facet of the ulna is a shallow concavity that occupies the entirety of the subrectangular proximal end of the bone. The simple geometry of the humeroulnar joint readily accommodates flexion/extension, protraction/retraction as well as rotation.

Joints distal to the ulna are formed by low convexities proximally

and shallow concavities distally, and thus permitted three degrees of freedom (flexion/extension, ab/adduction and rotation) (Fig. 6c, d). On the intermedium, ulnare and the largest of the proximal radials the distal articular surfaces extend onto the dorsal surfaces, indicating that extension was of predominant importance (Fig. 2a; see also Supplementary Information). The mobility of distal segments of the fin of *Tiktaalik* is further augmented by the transverse alignment of the joints distal to the epipodials (Figs 6 and 7). At least three transverse alignments are presently recognized: one at the joints distal to the radius, intermedium and ulnare; another at the joints distal to the next set of radials; and a third between the intermediate and distal radials. Additionally, numerous processes and rugosities, particularly on the ventral surfaces on the ulna, ulnare and distal radials, are suggestive of the presence of an extensive musculotendinous apparatus to control these fin segments (Fig. 2b). Not surprisingly, the enhanced mobility of a well-developed distal endoskeleton is accompanied by an apparent reduction in the length of the distal lepidotrichia (Supplementary Information).

Fin posture and function in *Tiktaalik*

The sister group of tetrapods²³ is now known to include fish with pectoral fins that can assume both fin-like and limb-like postures. In a quasi-planar position, with minimal flexion of the antebrachium,

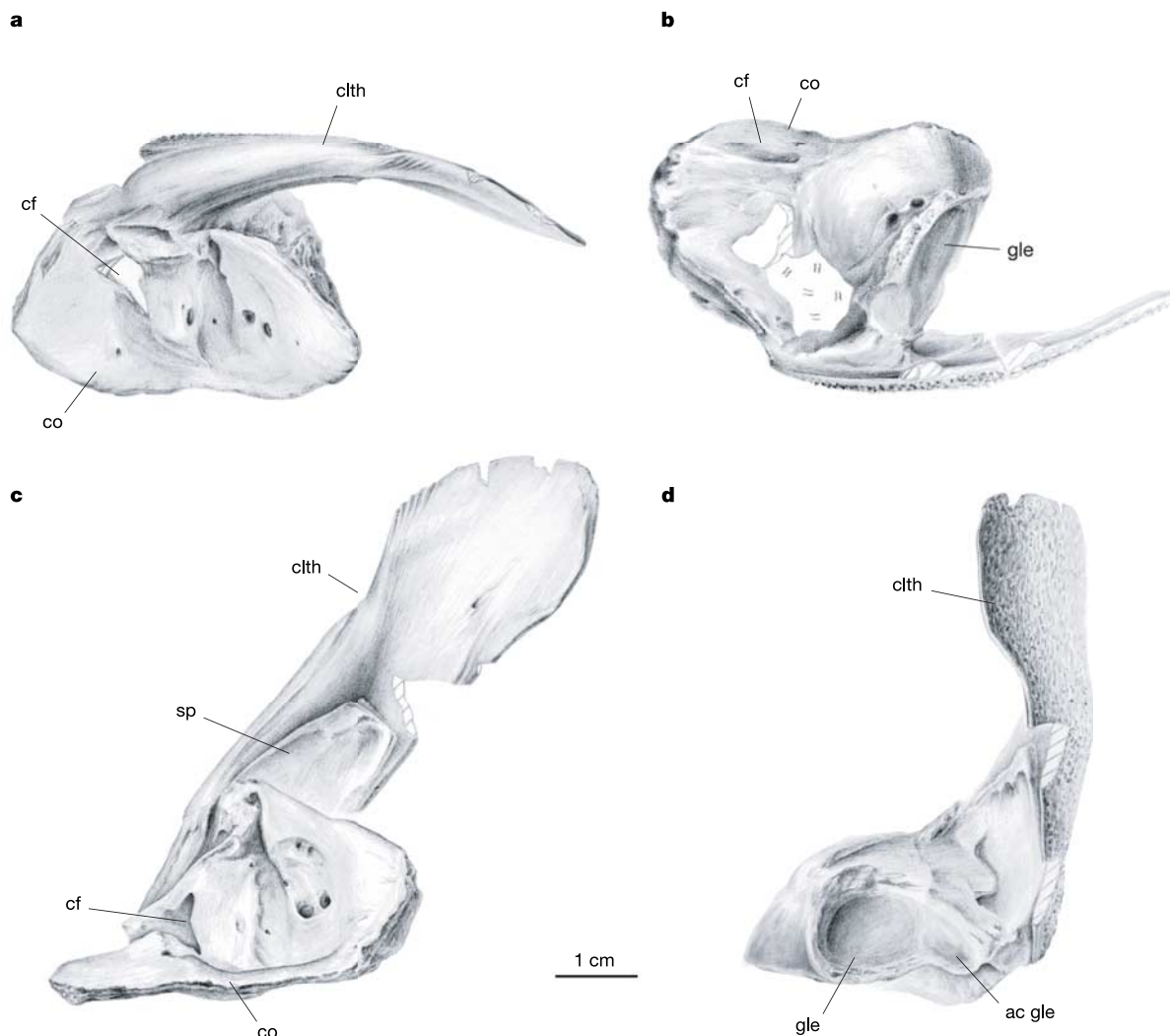


Figure 3 | Isolated right shoulder girdle of *Tiktaalik* (NUFV 112). **a**, Dorsal view; **b**, ventral view; **c**, medial view; **d**, posterior view. ac gle, anterior cam of glenoid facet; cf, coracoid foramen; clth, cleithrum; co, coracoid; gle, glenoid facet; sp, scapular process. The hatched area is covered by matrix.

the fin approximates a generalized posture for a sarcopterygian such as *Eusthenopteron*⁹ (Fig. 7a, c). But the fin could also assume a posture appropriate for a substrate-supported, upright stance by flexing the shoulder and elbow and extending the proximal and distal inter-radial joints (Fig. 7b, d). Multiple features enable the fin to prop the body in a limb-like manner: the base of the fin is positioned near the ventral surface of the body; glenohumeral architecture and transcoracoid musculature augment flexion and stability at the shoulder joint; a broad and deep posterior glenoid allows transmission of substantial propulsive stresses through the pectoral girdle; a robust coracoid plate provides broad areas for flexor muscle origins; elaborate ventral processes on the humerus represent extensive surface area for flexor insertions; flexion/extension, pronation/supination and rotation are possible at the elbow; and there is an expanded series of mobile proximal, intermediate and distal radials distal to the epipodials. Notably, the highly mobile yet robust distal fin segments could provide a stable but compliant extremity that could conform to complex and varied substrates.

The interpretation that the fins of *Tiktaalik* were used in supporting the body on a substrate is corroborated by the architecture of the axial skeleton. Expansion and imbrication of the ribs is a feature previously unknown in fish but seen in some early tetrapods such as *Ichthyostega*^{21,27}. The mechanical reinforcement of the spine engendered by costal overlap, together with a robust and mobile fin, suggest that both the axial and appendicular systems were playing a role in supporting the weight of the animal. With a

dorsoventrally compressed head and body, raised and dorsally placed eyes, and a mobile head that is independent of the shoulder girdle, *Tiktaalik* possesses a range of features consistent with locomotion on the water bottom, along the water margins, and on subaerial surfaces—an interpretation that is in accord with the shallow meandering stream deposits from which *Tiktaalik* was recovered²¹.

Tiktaalik and limb origins

The pectoral fins of *Tiktaalik* reveal that development of robusticity and mobility of the distal skeleton was underway before the origin of tetrapods. The array of joints in the distal fin is functionally similar to the multiple transverse joints that characterize the carpal, metacarpophalangeal and interphalangeal joints of the tetrapod manus. The distal endoskeleton of *Tiktaalik* invites direct comparisons to the wrists and digits of limbed vertebrates. The intermedium and ulnare of *Tiktaalik* have homologues to eponymous wrist bones of tetrapods with which they share similar positions and articular relations. In both *Tiktaalik* and early tetrapods, the ulnare is block-shaped and articulates with multiple radials or digits, whereas the intermedium is a simple rod. The formation of a mobile transverse joint at the distal margin of these bones in *Tiktaalik* presages the establishment of a functional proximal carpal joint. As in the digits and phalanges in a tetrapod limb, the inter-radial joints distal to this primordial wrist are more or less transversely aligned and capable of flexion and extension. The occurrence of multiple distally facing radial rows that

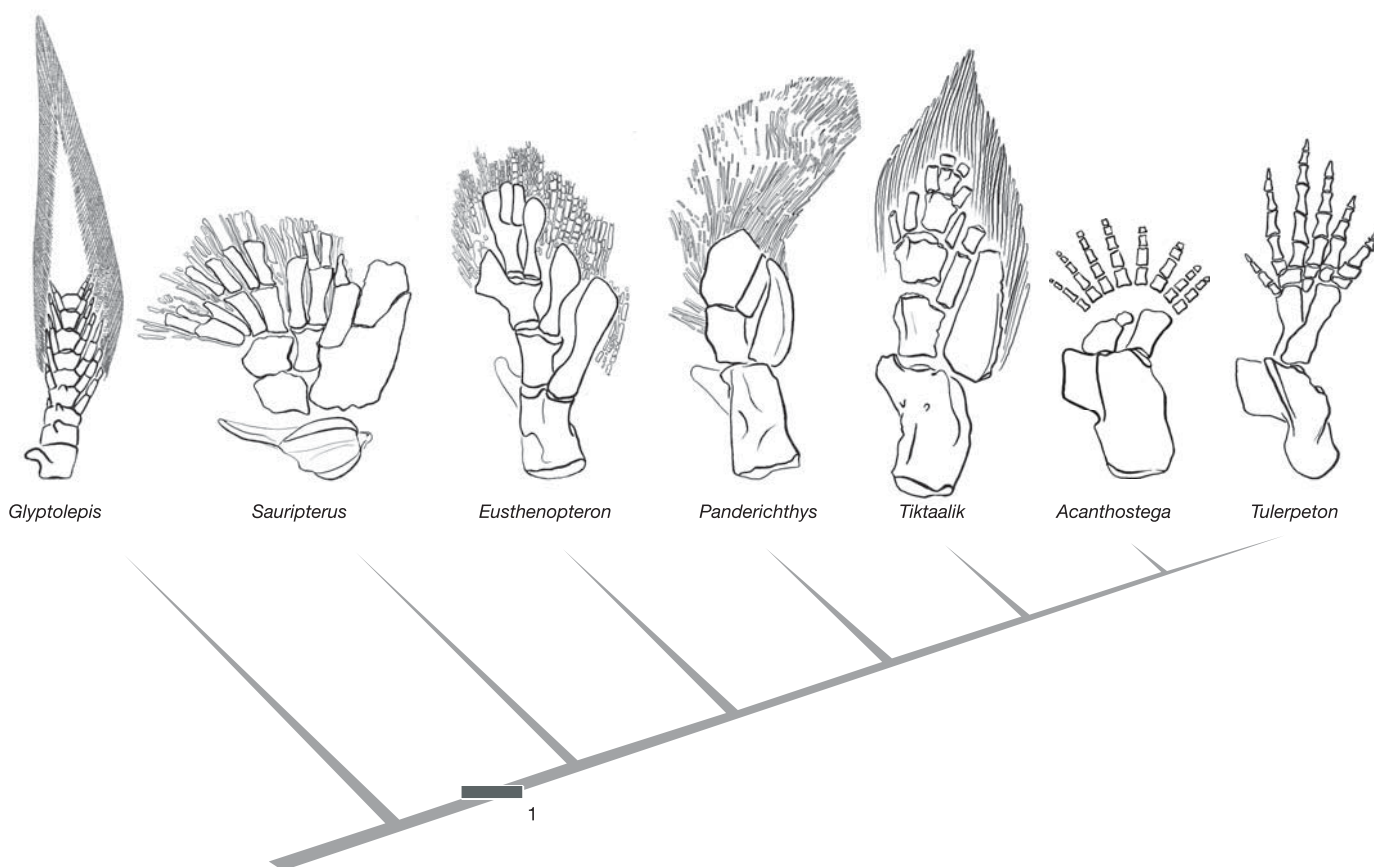


Figure 4 | Cladogram of the pectoral fins of taxa on the tetrapod stem.

Unlike other tetrapodomorph fishes (1), *Tiktaalik* has reduced the unjointed lepidotrichia, expanded the radials to a proximal, intermediate and distal series, and established multiple transverse joints in the distal fin. The fin also retains a mosaic of features seen in basal taxa. The central axis of enlarged endochondral bones is a pattern found in basal sarcopterygians and accords with hypotheses that a primitive fin axis is homologous to autopodial bones

of the tetrapod limb. In some features, *Tiktaalik* is similar to rhizodontids such as *Sauripterus*. These similarities, which are probably homoplastic, include the shape and number of radial articulations on the ulnare, the presence of extensive and branched endochondral radials, and the retention of unjointed lepidotrichia. Figures redrawn and modified from *Glyptolepis*³¹, *Eusthenopteron*⁹, *Panderichthys*³², *Acanthostega*²² and *Tulerpeton*³³.

are capable of flexion and extension is a likely antecedent condition to the dactyly of early tetrapods. The transformation of fins to limbs, then, probably entailed the elaboration and proliferation of structures, joints and functions already present in the fins of fish such as *Tiktaalik*.

The notion that the autopod is a developmental novelty of tetrapods may have been an artefact of relying exclusively on derived teleosts, such as zebrafish, in the analysis of limb origins^{1–6}. Zebrafish lack any fin bones homologous to tetrapod limbs. Therefore, a comparison of developmental mechanisms between zebrafish and tetrapods may be too coarse to provide insight into the developmental shifts that transformed fins—such as those in *Tiktaalik*—into the limbs of tetrapods. Accordingly, studies of the genetic basis of cartilage patterning in more basal actinopterygians or sarcopterygians may ultimately be more informative of the transformations that occurred in the Devonian period^{28–30}.

The presence of an elaborate tetrapod-like distal endoskeleton in rhizodontids and *Tiktaalik* may reflect the great antiquity of this feature or its parallel evolution in the two groups¹⁴. The kinematics of the endoskeleton, however, differ between these two taxa. Unlike

rhizodontids, such as *Sauripterus*¹⁴, the distal endoskeleton of both *Tiktaalik* and tetrapods contains multiple joints capable of extensive degrees of flexion and extension.

A fin axis that extends distal to the ulnare has been unknown in any tetrapodomorph⁷ until the discovery of *Tiktaalik*. As in porolepiforms and dipnoans, the axis of *Tiktaalik* lies in the centre of the fin. If the five radials of *Tiktaalik* are homologous to digital rays, then the axis of the tetrapod limb would extend from the humerus through digit three. Unfortunately, the absence of a well-defined axis in other tetrapodomorphs leaves uncertain whether a central axis is primitive for tetrapods or if it evolved separately in *Tiktaalik*. Testing these competing hypotheses awaits the discovery of other tetrapodomorph fins with axes that project into the distal fin.

The pectoral skeleton of *Tiktaalik* is transitional between fish fin and tetrapod limb. Comparison of the fin with those of related fish reveals that the manus is not a *de novo* novelty of tetrapods; rather, it was assembled in fishes over evolutionary time to meet the diverse challenges of life in the margins of Devonian aquatic ecosystems.

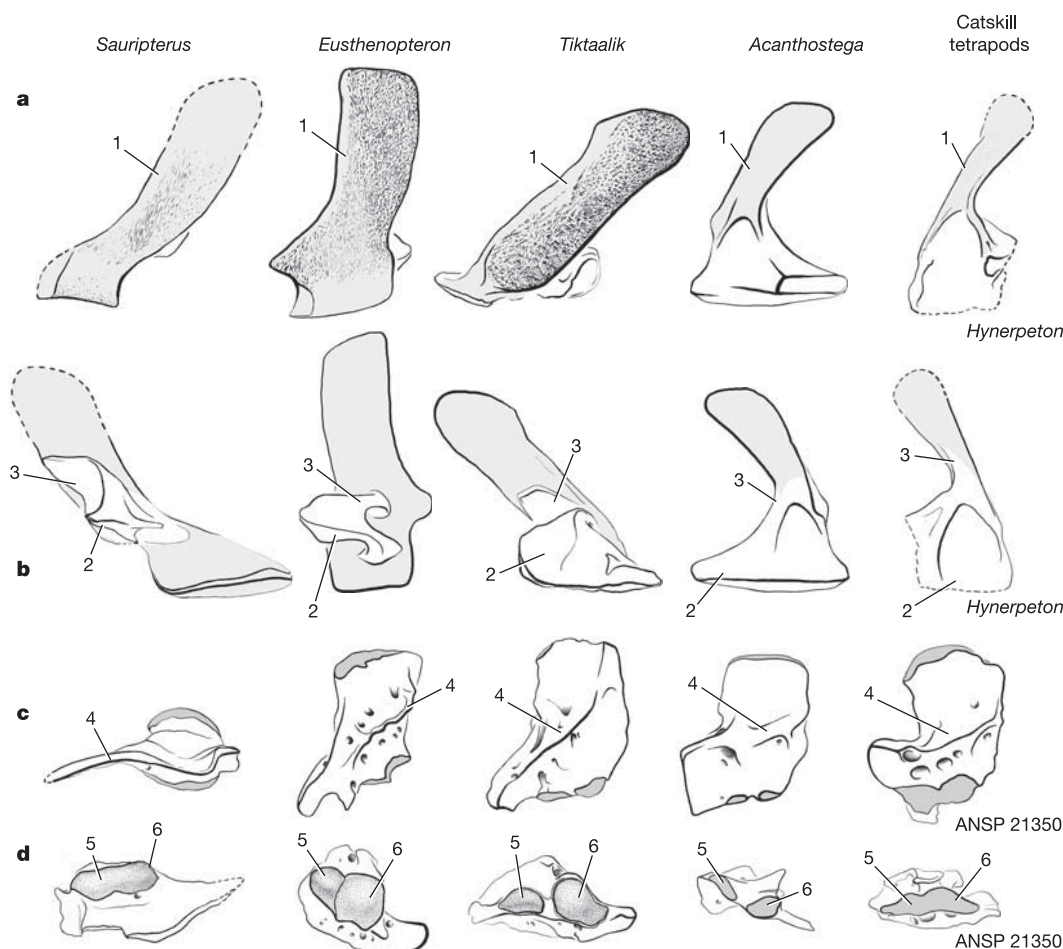


Figure 5 | A comparison of pectoral girdles and humeri in taxa along the tetrapod stem. **a, b**, Left pectoral girdles in lateral (**a**) and medial (**b**) views (dermal bone shaded). **c, d**, Left humeri in ventral view (**c**; articular surfaces shaded) and articular view (**d**; articular facets stippled; articular area shaded). *Tiktaalik* is intermediate between basal tetrapodomorphs and stem tetrapods in a wide range of features. **a**, As in basal taxa, the cleithrum (1) of *Tiktaalik* is heavily ornamented. *Tiktaalik* is intermediate in the degree to which the glenoid faces laterally, and hence, the appendage projects laterally. In both *Acanthostega* and *Tiktaalik* the appendage projects ventrolaterally from the body wall. **b**, *Tiktaalik* has an enlarged scapulocoracoid relative to

basal tetrapodomorphs that consists of an expanded coracoid plate (2) and a dorsally projecting scapular process (3). **c**, The humerus of *Tiktaalik* retains a diagonal ventral ridge (4) orientation as in *Eusthenopteron*, but has a prominent keel on the leading edge of the bone as in *Acanthostega* and ANSP 21350 (ref. 25). **d**, The radial (5) and ulnar (6) facets of *Tiktaalik* are separated by a narrow area of cortical bone and are dorsoventrally offset from one another. *Sauripterus* redrawn and modified from ref. 14; *Eusthenopteron* redrawn and modified from ref. 9; *Acanthostega* redrawn and modified from ref. 22.

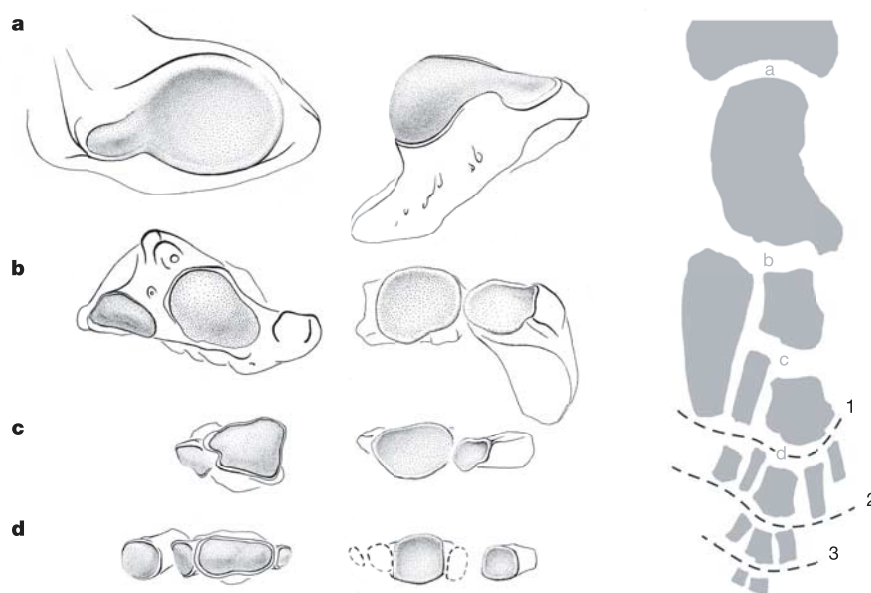


Figure 6 | Apposing joint surfaces of the left pectoral fin of NUV 109 in articular view. a, The shoulder joint consists of a large, shallow ball and socket posteriorly and a small accessory cam anteriorly. The cam serves to limit humeral protraction and supination. **b**, At the elbow joint, both epipodial facets of the humerus face slightly ventrally, with the radial facet offset anteroventrally from that for the ulna. **c**, The joints of the ulna and intermedium with the ulna are also offset from one another. As in the

anteroventral offset of the humeroradial joint, the joint of the intermedium is similarly offset relative to the joint for the ulnae. **d**, The distal ulnae, intermedium and radius form a transverse joint plane across the appendage; the ulnae and intermedium articulate with five radials at shallow concavoconvex joints. 1, 2 and 3 indicate proximodistal succession of joints that run transversely across the fin.

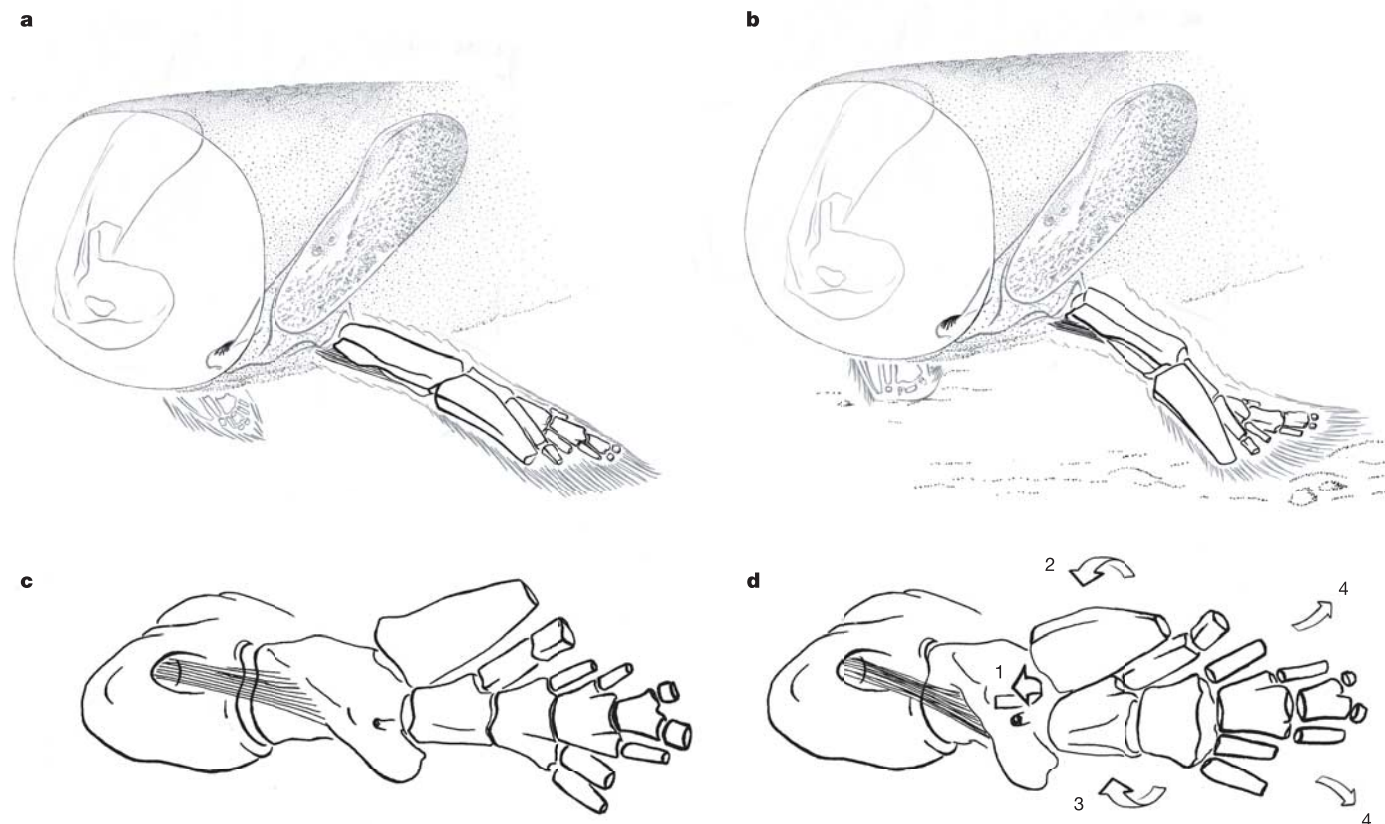


Figure 7 | Reconstruction of fin postures of *Tiktaalik*. a, b, Anterolateral view. **c, d**, Ventral view. **a, c**, Resting posture with the fin partially flexed at the antebrachium. In this position the radius is slightly more flexed than the ulna. **b, d**, Resistant contact with a firm substrate entails flexion at proximal joints and extension at distal ones. The shoulder joint is flexed by ventral

muscles, including the trans-coracoid muscle. The elbow is flexed (**d**, arrow 1), with slight pronation of the radius (**d**, arrow 2) and rotation of the ulna (**d**, arrow 3). The transverse joints distal to the ulnae and intermedium are extended (**d**, arrows 4).

Received 11 October 2005; accepted 8 February 2006.

1. Sordino, P. & Duboule, D. A molecular approach to the evolution of vertebrate paired appendages. *Trends Ecol. Evol.* **11**, 114–119 (1996).
2. Sordino, P., Hoeven, F. V. D. & Duboule, D. *Hox* gene expression in teleost fins and the origin of vertebrate digits. *Nature* **375**, 678–681 (1995).
3. Zakany, J., Fromental-Ramain, C., Warot, X. & Duboule, D. Regulation of number and size of digit by posterior *Hox* genes: a dose dependent mechanism with potential evolutionary implications. *Proc. Natl Acad. Sci. USA* **94**, 13695–13700 (1997).
4. Shubin, N., Tabin, C. & Carroll, S. Fossils, genes, and the evolution of animal limbs. *Nature* **388**, 639–648 (1997).
5. Capdevilla, J. & Izpisua-Belmonte, J. C. in *Perspectives on the Evolutionary Origin of Tetrapod Limbs. The Character Concept in Evolutionary Biology* (ed. Wagner, G. P.) 531–558 (Academic, San Diego, 2001).
6. Wagner, G. P. & Chiu, C.-H. The tetrapod limb: a hypothesis on its origin. *J. Exp. Zool. Mol. Dev. Evol.* **291**, 226–240 (2001).
7. Shubin, N. The evolution of paired fins and the origin of the tetrapod limb: phylogenetic and transformational approaches. *Evol. Biol.* **28**, 39–86 (1995).
8. Coates, M. I., Jeffery, J. E. & Ruta, M. Fins to limbs: what the fossils say. *Evol. Dev.* **4**, 390–401 (2002).
9. Andrews, S. M. & Westoll, T. S. The postcranial skeleton of *Eusteopteron foordi* Whiteaves. *Trans. R. Soc. Edinb.* **68**, 207–329 (1968).
10. Vorobyeva, E. I. The role of development and function in formation of “tetrapod-like” pectoral fins. *J. Comm. Biol.* **53**, 149–158 (1992).
11. Vorobyeva, E. I. & Kuznetsov, A. in *Fossil Fishes as Living Animals* (ed. Mark-Kurik, E.) 131–140 (Academy of Sciences of Estonia, Tallinn, 1992).
12. Daeschler, E. B. & Shubin, N. Fish with fingers? *Nature* **391**, 133 (1998).
13. Ahlberg, P. E. & Johanson, Z. Osteolepiforms and the ancestry of tetrapods. *Nature* **395**, 792–794 (1998).
14. Davis, M. C., Shubin, N. & Daeschler, E. B. A new specimen of *Sauripterus taylori* (Sarcopterygii; Osteichthyes) from the Famennian Catskill Formation of North America. *J. Vert. Paleontol.* **24**, 26–40 (2004).
15. Ahlberg, P. E., Luksevics, E. & Mark-Kurik, E. A near-tetrapod from the Baltic Middle Devonian. *Palaeontology* **43**, 533–548 (2000).
16. Cloutier, R. & Ahlberg, P. E. in *Interrelationships of Fishes* (eds Stiassny, M. L. J., Parenti, L. R. & Johnson, G. D.) 445–479 (Academic, New York, 1996).
17. Vorobyeva, E. I. Observations on two rhipidistian fishes from the Upper Devonian of Lode, Latvia. *Zool. J. Linn. Soc.* **70**, 191–201 (1980).
18. Vorobyeva, E. I. & Schultze, H.-P. in *Origins of the Higher Groups of Tetrapods* (eds Schultze, H.-P. & Trueb, L.) 68–109 (Cornell Univ. Press, Ithaca, 1991).
19. Schultze, H.-P. & Arsenault, M. The panderichthyid fish *Elpistostege*: a close relative of tetrapods? *Palaeontology* **28**, 292–309 (1985).
20. Schultze, H.-P. in *Devonian Fishes and Plants of Miguasha, Quebec, Canada* (eds Schultze, H.-P. & Cloutier, R.) 316–327 (Friedrich Pfeil, Munchen, 1996).
21. Daeschler, E. B., Shubin, N. H. & Jenkins, F. A. Jr. A Devonian tetrapod-like fish and the evolution of the tetrapod body plan. *Nature* doi:10.1038/nature04639 (this issue).
22. Coates, M. I. The Devonian tetrapod *Acanthostega gunnari* Jarvik: postcranial anatomy, basal tetrapod interrelationships and patterns of skeletal evolution. *Trans. R. Soc. Edinb. Earth Sci.* **87**, 363–421 (1996).
23. Daeschler, E. B., Shubin, N. H., Thomson, K. S. & Amaral, W. W. A Devonian tetrapod from North America. *Science* **265**, 639–642 (1994).
24. Vorobyeva, E. I. The shoulder girdle of *Panderichthys rhombolepis* (Gross) (Crossopterygii), Upper Devonian, Latvia. *Geobios.* **19**, 285–288 (1995).
25. Shubin, N. H., Daeschler, E. B. & Coates, M. I. The early evolution of the tetrapod humerus. *Science* **304**, 90–93 (2004).
26. Shubin, N. H. & Alberch, P. A. Morphogenetic approach to the origin and basic organization of the tetrapod limb. *Evol. Biol.* **20**, 319–387 (1986).
27. Ahlberg, P. E., Clack, J. A. & Blom, H. The axial skeleton of the Devonian tetrapod *Ichthyostega*. *Nature* **437**, 137–140 (2005).
28. Davis, M. C., Shubin, N. H. & Force, A. Pectoral fin and girdle development in the basal Actinopterygians *Polyodon spathula* and *Acipenser transmontanus*. *J. Morphol.* **262**, 608–628 (2004).
29. Mabee, P. M. & Noordsy, M. Development of the paired fins in the paddlefish, *Polyodon spathula*. *J. Morphol.* **261**, 334–344 (2004).
30. Metscher, B. D. et al. Expression of *Hoxa-11* and *Hoxa-13* in the pectoral fin of a basal ray-finned fish, *Polyodon spathula*: implications for the origin of tetrapod limbs. *Evol. Dev.* **7**, 186–195 (2005).
31. Ahlberg, P. E. Paired fin skeletons and relationships of the fossil group Porolepiformes (Osteichthyes: Sarcopterygii). *Zool. J. Linn. Soc.* **96**, 119–166 (1989).
32. Vorobyeva, E. I. Morphology of the humerus in the Rhipidistian Crossopterygii and the origin of tetrapods. *Paleontol. J.* **34**, 632–641 (2000).
33. Lebedev, O. A. & Coates, M. I. The postcranial skeletal of the Devonian tetrapod *Tulerpeton curtum* Lebedev. *Zool. J. Linn. Soc.* **114**, 307–348 (1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The illustrations are the work of K. Monoyios. Specimen preparation was performed by C. F. Mullison and B. Masek. NUFV 108 was discovered and quarried by S. Gatesy. Permits to conduct this research were granted by the Nunavut Ministry of Culture, Languages, Elders and Youth (D. Stenton and J. Ross), the Grise Fiord Hamlet, and the Iviq Hunters and Trappers Association. Field operations received logistical support from the Polar Continental Shelf Project (B. Hycyk, D. Maloley, J. MacEachern, W. Benoit, G. Benoit, N. Couture, H. Gordon, D. Clouthier and D. Mueller) and collections support from the Canadian Museum of Nature (S. Cumbaa and K. Shepherd). The elders of the Inuit Qaujimajatuqangit Katimajit of the Government of Nunavut provided the generic name for the new taxon. M. Shubin coordinated the naming effort. A. Embry and U. Mayr provided guidance at the inception of the field project. M. Coates commented on a draft of the manuscript. Field assistance (1999–2004) was provided by W. Amaral, B. Atagootak, J. Conrad, M. Davis, J. Downs, S. Gatesy, S. Madsen, K. Middleton, K. Monoyios, C. Schaff, M. Shapiro, R. Shearman and C. Sullivan. This research was supported by a patron of our research, the Academy of Natural Sciences, the Putnam Expeditionary Fund (Harvard University), the University of Chicago, the National Science Foundation and the National Geographic Society Committee for Research and Exploration.

Author Contributions N.H.S. and E.B.D. conceived and co-directed the project. F.A.J. Jr collaborated on all phases of the research.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to N.H.S. (nshubin@uchicago.edu).

LETTERS

A debris disk around an isolated young neutron star

Zhongxiang Wang¹, Deepto Chakrabarty¹ & David L. Kaplan¹

Pulsars are rotating, magnetized neutron stars that are born in supernova explosions following the collapse of the cores of massive stars. If some of the explosion ejecta fails to escape, it may fall back onto the neutron star¹ or it may possess sufficient angular momentum to form a disk². Such ‘fallback’ is both a general prediction of current supernova models³ and, if the material pushes the neutron star over its stability limit, a possible mode of black hole formation⁴. Fallback disks could dramatically affect the early evolution of pulsars^{2,5}, yet there are few observational constraints on whether significant fallback occurs or even the actual existence of such disks. Here we report the discovery of mid-infrared emission from a cool disk around an isolated young X-ray pulsar. The disk does not power the pulsar’s X-ray emission but is passively illuminated by these X-rays. The estimated mass of the disk is of the order of 10 Earth masses, and its lifetime ($\geq 10^6$ years) significantly exceeds the spin-down age of the pulsar, supporting a supernova fallback origin. The disk resembles protoplanetary disks seen around ordinary young stars⁶, suggesting the possibility of planet formation around young neutron stars.

The so-called anomalous X-ray pulsars⁷ (AXPs) are a group of young ($\leq 10^5$ yr) neutron stars with spin periods falling in a narrow range (5–12 s), no evidence for binary companions, and whose X-ray luminosity ($\sim 10^{36}$ erg s⁻¹) greatly exceeds their rate of rotational energy loss ($\sim 10^{33}$ erg s⁻¹). AXPs are generally believed to be ‘magnetars’⁸, which are isolated neutron stars with exceptionally strong ($\sim 10^{14}$ G) surface magnetic field strengths and whose magnetic energy ultimately powers their X-ray emission. An alternative

explanation for AXPs attributes their X-ray emission to accretion from a residual debris disk^{9–11}, but this model has had difficulties explaining observations¹². The brightest known AXP is the 8.7-s pulsar 4U 0142+61, at a distance of 3.9 kpc (ref. 13). Besides its X-ray emission, the pulsar also has known optical¹⁴ and near-infrared¹² (near-IR) counterparts.

As part of a systematic search for fallback disks around young neutron stars, we observed the field around 4U 0142+61 in the 4.5 μ m and 8.0 μ m bands with the Infrared Array Camera (IRAC) on the Spitzer Space Telescope to look for the IR excess predicted by models for an X-ray heated fallback disk¹⁵. We found a candidate mid-IR counterpart at the pulsar’s position in both bands (Fig. 1) that has very unusual IR colours (Fig. 2). On the basis of the position coincidence and colours, we conclude that we have identified the mid-IR counterpart of 4U 0142+61.

We can reconstruct the observed low-energy spectral energy distribution of 4U 0142+61 by combining our Spitzer data with existing IR and optical data^{12,16} (Fig. 3). We may then infer the intrinsic spectrum by correcting for interstellar reddening. Although this reddening correction has little effect on the mid-IR data, it significantly affects the optical data. For any reasonable choice of reddening ($2.6 \leq A_V \leq 5.1$; ref. 12), the optical and IR data clearly arise from two different spectral components. The optical (VRI) emission is consistent with a power-law spectrum¹², is pulsed at the spin period¹⁷, and is presumably of magnetospheric origin; we will not consider the optical component further, but will instead confine ourselves to the distinct IR component. Hereafter, we adopt the most likely reddening value of $A_V = 3.5$ (ref. 13).

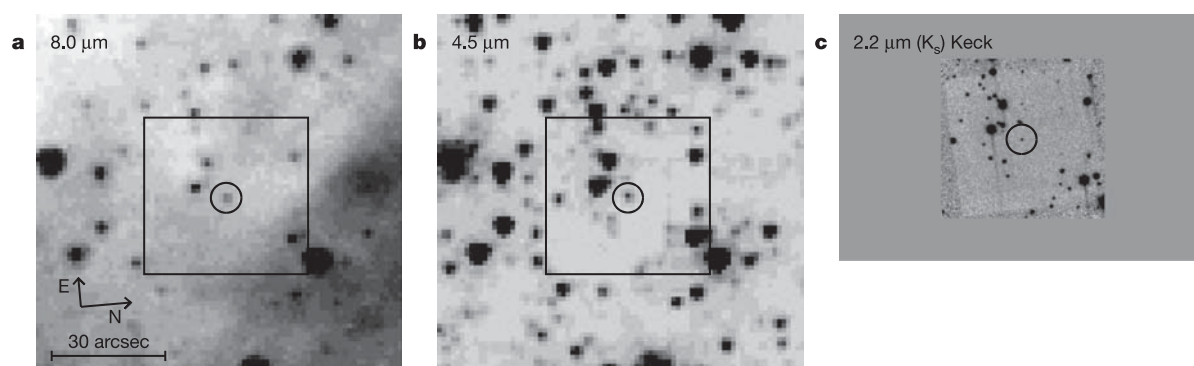


Figure 1 | Infrared images of the 4U 0142+61 field. **a, b**, Spitzer/IRAC mid-IR images taken on 17 January 2005 in the 8.0 μ m band (**a**; 75 min exposure) and the 4.5 μ m band (**b**; 77 min exposure). A 0.15 arcsec astrometric solution was derived using 14 unblended field stars from the 2MASS IR point source catalogue. A candidate pulsar counterpart (indicated by a 4-arcsec-radius circle) is detected in both IRAC images at a position of right ascension 01 h 46 min 22.40 s, declination $+61^\circ 45' 02.9''$ (equinox J2000.0), with a 1 σ confidence radius of 0.7 arcsec. This is consistent with the known pulsar position¹⁴, lying within 0.3 arcsec. Based

on the source number density in the 4.5 μ m image, the chance coincidence probability is only 0.2%. The source fluxes were $51.9 \pm 5.2 \mu$ Jy (8.0 μ m) and $36.3 \pm 3.6 \mu$ Jy (4.5 μ m), where 1μ Jy = 10^{-29} erg cm⁻² s⁻¹ Hz⁻¹. **c**, A deep K_s-band (2.2 μ m) near-IR comparison image of the central part of the IRAC field, taken on 1 November 2001 from the Keck-I telescope in Hawaii. (The corresponding subfield is indicated by a box in **a** and **b**.) Besides the pulsar itself ($K_s \approx 20$), there are no other objects detected within 4.0 arcsec of the source position down to a magnitude limit of $K_s \approx 22$. The scale and orientation for **b** and **c** are the same as shown in **a**.

¹Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

We begin by noting that a non-thermal (self-absorbed synchrotron) origin for the IR component can be ruled out, because associating the turnover near $4.5\text{ }\mu\text{m}$ with a synchrotron self-absorption break requires an emission region with an implausibly weak ($\sim 10^4\text{ G}$) magnetic field strength within a small ($\sim 100\text{ km}$) radius of an active, highly magnetized pulsar. Similarly, the shape of the IR spectrum suggests a thermal origin, but the simplest case of a single-temperature (920 K) blackbody gives a mediocre fit to the data. Moreover, the only plausible blackbody source would be a large planet or extremely low-mass stellar companion, but the implied blackbody emission radius ($>5R_\odot$, where $R_\odot$ is the solar radius) is much too large.

A multi-temperature (700–1,200 K) thermal model fits the data better, and thus naturally suggests an extended disk or shell origin for the IR emission. We reject a shell geometry as it would produce $\sim 1,000$ times the X-ray and optical absorption observed towards the AXP. As for a disk, it has already been shown that an accretion disk model (where the pulsar's X-ray flux is entirely powered by disk accretion onto the neutron star) is ruled out, as it greatly overpredicts the optical flux¹⁴. Instead, we consider a passive disk illuminated by the X-ray pulsar. At these low temperatures, the disk's continuum emission would be entirely from dust. Indeed, we note that our observed IR spectrum is quite similar to those of dusty protostellar disks⁶, in which the disk is passively illuminated by the finite-radius central source. In our case, the neutron star is effectively an X-ray point source, but assuming there is enough gas to provide pressure support for the dust, we would expect any surrounding disk to be slightly flared (with corresponding disk thickness $h \propto r^{9/7}$ at disk radius r ; ref. 18) and thus subject to X-ray irradiation.

As additional support for an X-ray-heated disk model for the IR flux, we note that four of the eight known AXPs have known near-IR counterparts (including 4U 0142+61), and all of these have similar IR/X-ray flux ratios, of the order of 10^{-4} (ref. 19). In at least one source, there is even a detected correlation between the near-IR and

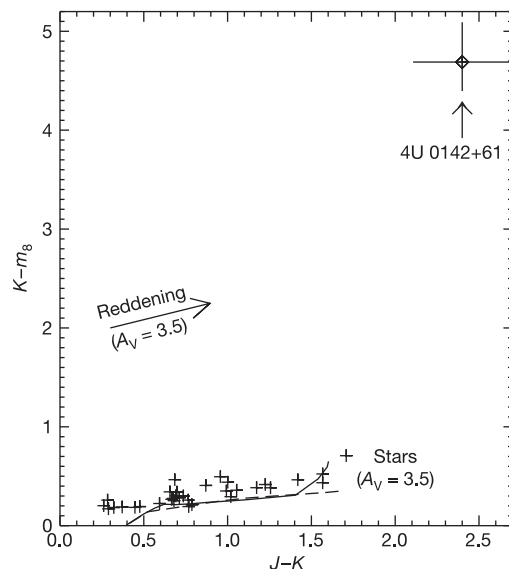


Figure 2 | Infrared colour-colour diagram for 4U 0142+61 (diamond) and 33 field stars (small crosses) from the 2MASS point source catalogue. We converted the $8\text{ }\mu\text{m}$ fluxes to magnitudes (m_8) using a zero point of 63.1 Jy, and used published values of the near-IR magnitudes for the pulsar¹⁶. ($K_s = 19.9$ is an average value for the pulsar's variable K_s -band flux¹².) The field stars have colours consistent with those of reddened ($A_V = 3.5$) main sequence stars (solid line) or red giants (dashed line). An example $A_V = 3.5$ reddening vector is also plotted. The object at the position of 4U 0142+61 has highly unusual colours (independent of reddening) and a small chance coincidence probability, and we conclude that it is the pulsar counterpart. The error bars for 4U 0142+61 are 1 σ .

X-ray fluxes²⁰. This may indicate that the IR emission from all AXPs arises in an X-ray-heated debris disk.

We fit the power-law-subtracted IR spectrum from 4U 0142+61 with an X-ray-heated disk model¹⁸ with five parameters: source distance d , disk inclination angle i , inner radius r_{in} , outer radius r_{out} , and X-ray albedo η_d . This model predicts that the disk's effective temperature T at radius r is

$$T(r) \approx 5,030\text{ K} (1 - \eta_d)^{2/7} \left(\frac{d}{3.9\text{ kpc}} \right)^{4/7} \left(\frac{r}{R_\odot} \right)^{-3/7} \quad (1)$$

where an unabsorbed 0.5–10 keV X-ray flux of $4.8 \times 10^{-10}\text{ erg cm}^{-2}\text{ s}^{-1}$ from the pulsar (ref. 12) is assumed to illuminate the disk. The total disk flux emitted at a given frequency is found by integrating the thermal emission over the entire projected ($\cos i$) disk surface between r_{in} and r_{out} . For our fits, we fixed $d = 3.9\text{ kpc}$ and $\cos i = 0.5$ and assumed $\eta_d > 0.9$, as favoured empirically by observations of X-ray-heated neutron star disks in low-mass X-ray binaries²¹. We obtained a good fit with $r_{\text{in}} = 2.9R_\odot$, $r_{\text{out}} = 9.7R_\odot$, and $\eta_d = 0.97$.

The inner disk temperature is well constrained to be $T(r_{\text{in}}) \approx 1,200\text{ K}$ by the shape of the near-IR spectrum. This temperature is comparable to the sublimation temperature of dust (1,000–2,000 K, depending upon grain composition), suggesting that r_{in} may be set by the X-ray destruction of dust. Disk gas could persist inside r_{in} without any continuum emission at these temperatures, cooling instead by line emission. Alternatively, r_{in} might also have been set by interactions with the pulsar magnetosphere at some time in the past, which is consistent with the fact that the pulsar's

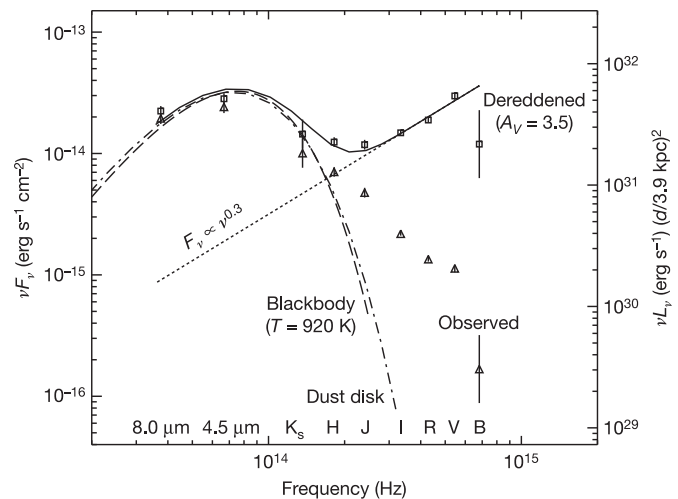


Figure 3 | Optical/infrared spectral energy distribution of 4U 0142+61.

The vertical axes are both scaled by frequency ν . The left axis shows ν -scaled flux per unit frequency, νF_ν ; the right axis shows ν -scaled luminosity per unit frequency, νL_ν , for an assumed distance of $d = 3.9\text{ kpc}$. The K_s point is an average value for the pulsar's variable emission in that band¹². The triangles indicate the observed optical/IR flux while the squares indicate the dereddened flux assuming $A_V = 3.5$. The dereddened optical VRI flux is well fitted by a $F_\nu \propto \nu^{0.3}$ power-law component (short-dashed line) up to the B-band ($0.44\text{ }\mu\text{m}$) break and is relatively constant in contrast to the near-IR variability (0.5 mag). This power law is presumably magnetospheric in origin, an assertion supported by the optical pulsations that cannot be due to X-ray reprocessing¹⁷. The power-law-subtracted IR spectrum is poorly fitted by a simple blackbody model (long-dashed curve), probably excluding the presence of large planets or a very-low-mass stellar companion. The best-fit passive debris disk model, including irradiation by the central X-ray source, is plotted as the dot-dashed curve. The sum (solid curve) of the disk and power-law components fits the optical/IR data (below the break) well. An accretion disk (including irradiation by accretion-powered X-rays from the pulsar) can be ruled out because it significantly overpredicts the optical flux¹⁴.

light cylinder radius ($r_{\text{LC}} = 0.61R_{\odot}$) lies well inside r_{in} . The fit value for r_{in} is sensitive to our choice of η_{d} ; for $\eta_{\text{d}} = 0.5$, the disk is further from the pulsar with $r_{\text{in}} \approx 19R_{\odot}$, although the fit quality is poorer.

By contrast, the uncertainty in r_{out} is dominated by our ignorance of the spectrum at wavelengths longer than $8.0\text{ }\mu\text{m}$, as the location of r_{out} is inferred from the frequency below which the spectrum turns over to the Rayleigh–Jeans ($F_{\nu} \propto \nu^2$) regime. (Here F_{ν} is the flux per unit frequency at frequency ν .) As this is poorly constrained by our data, our fitted value for r_{out} should be regarded as a lower limit. Observations in the far-IR or millimetre (MM) bands will be required to determine r_{out} more precisely.

Another benefit of far-IR/MM observations is that dust emission in these bands is typically optically thin, allowing for a direct measure of the disk mass⁶. In the absence of far-IR data, we can at least assume that the flux at 1 mm does not exceed our $8.0\text{-}\mu\text{m}$ flux as the spectrum appears to be falling at wavelengths longer than $4.5\text{ }\mu\text{m}$. This sets an upper limit on the total disk mass (both dust and gas) of⁶

$$M_{\text{d}} < 3 \times 10^{-3} M_{\odot} \left(\frac{F_{\text{MM}}}{50 \mu\text{Jy}} \right) \left(\frac{d}{3.9 \text{ kpc}} \right)^2 \times \left(\frac{T(r_{\text{out}})}{300 \text{ K}} \right)^{-1} \left(\frac{\kappa_{\text{MM}}}{0.01 \text{ cm}^2 \text{ g}^{-1}} \right)^{-1} \quad (2)$$

where $M_{\odot}$ is the solar mass, F_{MM} and κ_{MM} are respectively the flux and opacity at 1 mm , and we have assumed a dust-to-gas ratio of 1% in the disk. (Note that our assumptions for κ_{MM} and the dust-to-gas ratio do not account for the high metallicity and unusual composition possible for supernova-processed material.) In protostellar disks, F_{MM} is generally around two orders of magnitude smaller than the mid-IR flux⁶, and this is consistent with an extrapolation of our best-fit disk model for 4U 0142+61. This suggests a disk mass of the order of $M_{\text{d}} \approx 10^{-5} M_{\odot}$, comparable to the total mass of Earth-mass planets detected around the old millisecond radio pulsar PSR B1257+12 (refs 22, 23).

In order to consider the lifetime of the disk, τ_{d} , we first need an estimate of the disk's mass loss rate. For an upper limit, we can interpret the pulsar's spin period derivative²⁴ ($\dot{P} = 2 \times 10^{-12}$) in terms of magnetic 'propeller' torques²⁵ to argue that $-\dot{M}_{\text{d}} \leq 10^{-11} M_{\odot} \text{ yr}^{-1}$. (This presumes that the disk penetrates inside r_{in} and into the pulsar light cylinder to interact with the magnetosphere at the corotation radius, $r_{\text{co}} = 0.01R_{\odot}$.) The resulting disk lifetime $\tau_{\text{d}} = M_{\text{d}}/|\dot{M}_{\text{d}}|$ is much longer than the pulsar's spin-down age, $\tau_{\text{p}} = P/2\dot{P} \approx 10^5 \text{ yr}$, consistent with a supernova fallback origin. This also suggests that the disk did not start out with considerably more mass than it currently contains.

Our detection of a fallback disk is, to our knowledge, the first direct evidence for supernova fallback in any context, and it thus bolsters the notion of fallback in a variety of other circumstances. For example, direct fallback (without a disk) in the seconds to hours following core collapse has been proposed as a mechanism for black hole formation following a supernova explosion⁴, and may explain the absence of a detectable pulsar in the nearby supernova SN 1987A. However, if a fallback disk forms around a new neutron star, there are several mechanisms that can limit the disk lifetime^{26,27}, including accretion, magnetic propeller expulsion, and ablation by a pulsar wind, with the details depending upon the spin rate and magnetic moment of the neutron star. In addition, giant X-ray flares such as those seen in some magnetars²⁸ may also severely limit the lifetime of fallback disks. Our discovery implies that it is possible for a disk to survive as long as 10^5 yr , although 4U 0142+61 may be an exceptional case.

If fallback disks are common around all young neutron stars, then pulsar spindown may be caused by disk–magnetosphere interactions in addition to magnetic dipole torques, rendering traditional estimates of pulsar ages and magnetic moments unreliable^{2,5}. On the other hand, if fallback disks are peculiar to AXPs, this may point to a

distinct formation mechanism for magnetars. It may also explain the narrow clustering of spin periods for AXPs^{11,12}, despite the need for magnetar activity to account for their X-ray flux. However, it is difficult to discriminate among these possibilities because there are few constraints on the presence of disks around young pulsars. Previous searches, motivated by the discovery of a pulsar planetary system²², have concentrated on middle-aged and old radio pulsars with no success²⁹. Our detection in 4U 0142+61 is probably due to a combination of the X-ray heating of the disk and the relative youth of the pulsar (compared to τ_{d}). We note that energetic particles in a radio pulsar wind may be as effective as X-rays at heating a disk, so that both X-ray and radio pulsars are promising targets, and additional debris disks may yet be detected in a suitably selected sample. However, the strong non-thermal emission from a young pulsar or supernova remnant may mask a disk signature in some cases.

By analogy with protoplanetary disks around ordinary young stars⁶, the presence of a debris disk around a neutron star naturally raises the possibility of planet formation. A debris disk is the presumed origin^{26,23} of the system of Earth-mass planets already known to exist around the old ($\sim 10^9 \text{ yr}$) millisecond pulsar PSR B1257+12, which was detected through pulsar timing residuals^{22,23}. However, millisecond pulsars are generally believed to have been spun up through sustained disk accretion, and the formation and survival of a debris disk and planetary system in this context are difficult to explain²⁷. Our discovery does not directly address this system, but it does demonstrate that a debris disk around a neutron star is possible. Despite this, it is not clear whether the radiation-rich environment around 4U 0142+61 is suitable for the formation and survival of planets²⁷. On the other hand, as planet formation probably requires $\sim 10^6 \text{ yr}$ after the disk's formation³⁰, magnetar activity may well have subsided by that time. In any case, searches for planets around young pulsars are likely to be difficult. Pulsar timing searches for planets will have poor sensitivity in AXPs owing to their slow spin periods, while timing noise in the faster-spinning young radio pulsars will limit sensitivity to planets in these systems.

Received 5 August 2005; accepted 21 February 2006.

- Colgate, S. A. Neutron-star formation, thermonuclear supernovae, and heavy-element reimplosion. *Astrophys. J.* **163**, 221–230 (1971).
- Michel, F. C. & Dessler, A. J. Pulsar disk systems. *Astrophys. J.* **251**, 654–664 (1981).
- Heger, A., Fryer, C. L., Woosley, S. E., Langer, N. & Hartmann, D. H. How massive single stars end their life. *Astrophys. J.* **591**, 288–300 (2003).
- Chevalier, R. A. Neutron star accretion in a supernova. *Astrophys. J.* **346**, 847–859 (1989).
- Menou, K., Perna, R. & Hernquist, L. Disk-assisted spin-down of young radio pulsars. *Astrophys. J.* **554**, L63–L66 (2001).
- Beckwith, S. V. W., Sargent, A. I., Chini, R. S. & Guesten, R. A survey for circumstellar disks around young stellar objects. *Astron. J.* **99**, 924–945 (1990).
- Woods, P. & Thompson, C. in *Compact Stellar X-Ray Sources* (eds Lewin, W. & van der Klis, M.) (Cambridge Univ. Press, in the press); preprint at (<http://arXiv.org/astro-ph/0406133>) (2006).
- Thompson, C. & Duncan, R. The soft gamma repeaters as very strongly magnetized neutron stars. II. Quiescent neutrino, X-ray, and Alfvén wave emission. *Astrophys. J.* **473**, 322–342 (1996).
- van Paradijs, J., Taam, R. E. & van den Heuvel, E. P. J. On the nature of the 'anomalous' 6-s X-ray pulsars. *Astron. Astrophys.* **299**, L41–L44 (1995).
- Alpar, M. A. On young neutron stars as propellers and accretors with conventional magnetic fields. *Astrophys. J.* **554**, 1245–1254 (2001).
- Chatterjee, P., Hernquist, L. & Narayan, R. An accretion model for anomalous X-ray pulsars. *Astrophys. J.* **534**, 373–379 (2000).
- Hulleman, F., van Kerkwijk, M. H. & Kulkarni, S. R. The anomalous X-ray pulsar 4U 0142+61: variability in the infrared and a spectral break in the optical. *Astron. Astrophys.* **416**, 1037–1045 (2004).
- Durant, M. & van Kerkwijk, M. H. The red clump method for reddening and distance determination for the anomalous X-ray pulsars. *Astrophys. J.* (submitted).
- Hulleman, F., van Kerkwijk, M. H. & Kulkarni, S. R. An optical counterpart to the anomalous X-ray pulsar 4U0142+61. *Nature* **408**, 689–692 (2000).
- Perna, R., Hernquist, L. & Narayan, R. Emission spectra of fallback disks around young neutron stars. *Astrophys. J.* **541**, 344–350 (2000).

16. Israel, G., *et al.* in *Young Neutron Stars and Their Environments* (eds Camilo, F. & Gaensler, B. M.) 247–250 (IAU Symp. 218, Astron. Soc. Pacific, San Francisco, 2004).
17. Kern, B. & Martin, C. Optical pulsations from the anomalous X-ray pulsar 4U0142+61. *Nature* **417**, 527–529 (2002).
18. Vrtiljek, S. D. *et al.* Observations of Cygnus X-2 with IUE—ultraviolet results from a multiwavelength campaign. *Astron. Astrophys.* **235**, 162–173 (1990).
19. Durant, M. & van Kerkwijk, M. H. The broadband spectrum and infrared variability of the magnetar AXP 1E 1048.1–5937. *Astrophys. J.* **627**, 376–382 (2005).
20. Tam, C. R., Kaspi, V. M., van Kerkwijk, M. H. & Durant, M. Correlated infrared and X-ray flux changes following the 2002 outburst of the anomalous X-ray pulsar 1E 2259+586. *Astrophys. J.* **617**, L53–L56 (2004).
21. de Jong, J. A., van Paradijs, J. & Augusteijn, T. Reprocessing of X-rays in low-mass X-ray binaries. *Astron. Astrophys.* **314**, 484–490 (1996).
22. Wolszczan, A. & Frail, D. A. A planetary system around the millisecond pulsar PSR1257+12. *Nature* **355**, 145–147 (1992).
23. Wolszczan, A. Confirmation of Earth-mass planets orbiting the millisecond pulsar PSR B1257+12. *Science* **264**, 538–542 (1994).
24. Gavril, F. P. & Kaspi, V. M. Long-term Rossi X-Ray Timing Explorer monitoring of anomalous X-ray pulsars. *Astrophys. J.* **567**, 1067–1076 (2002).
25. Ekşi, K. Y., Hernquist, L. & Narayan, R. Where are all the fallback disks? Constraints on propeller systems. *Astrophys. J.* **623**, L41–L44 (2005).
26. Phinney, E. S. & Hansen, B. M. S. in *Planets Around Pulsars* (eds Phillips, J. A., Thorsett, S. E. & Kulkarni, S. R.) 371–390 (ASP Conf. Vol. 36, Astron. Soc. Pacific, San Francisco, 1993).
27. Miller, M. C. & Hamilton, D. P. Implications of the PSR 1257+12 planetary system for isolated millisecond pulsars. *Astrophys. J.* **550**, 863–870 (2001).
28. Hurley, K. *et al.* An exceptionally bright flare from SGR 1806–20 and the origins of short-duration γ -ray bursts. *Nature* **434**, 1098–1103 (2005).
29. Löhmer, O., Wolszczan, A. & Wielebinski, R. A search for cold dust around neutron stars. *Astron. Astrophys.* **425**, 763–766 (2004).
30. Lin, D. N. C., Woosley, S. E. & Bodenheimer, P. H. Formation of a planet orbiting pulsar 1829–10 from the debris of a supernova explosion. *Nature* **353**, 827–829 (1991).

Acknowledgements We thank M. van Kerkwijk for sharing the K_s -band image of 4U 0142+61. We also thank A. Alpar, L. Bildsten, E. Chiang, M. Durant, E. Dwek, Y. Ekşi, L. Hernquist, M. Jura and R. Narayan for discussions. This work is based on observations made with the Spitzer Space Telescope, which is operated by JPL/Caltech under a NASA contract. Support for this work was provided by NASA through a contract issued by JPL/Caltech. DLK was supported by a Pappalardo Fellowship.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.C. (deepto@space.mit.edu).

LETTERS

Isotopic enhancements of ^{17}O and ^{18}O from solar wind particles in the lunar regolith

Trevor R. Ireland¹, Peter Holden¹, Marc D. Norman^{1,2} & Jodi Clarke^{1†}

Differences in isotopic abundances between meteorites and rocks on Earth leave unclear the true composition of the gas out of which the Solar System formed^{1–4}. The Sun should have preserved in its outer layers the original composition, and recent work has indicated that the solar wind is enriched in ^{16}O , relative to Earth, Mars and bulk meteorites⁵. This suggests that self-shielding of CO due to photo-dissociation, which is a well understood process in molecular clouds, also led to evolution in the isotopic abundances in the early Solar System. Here we report measurements of oxygen isotopic abundances in lunar grains that were recently exposed to the solar wind. We find that ^{16}O is under-abundant, opposite to an earlier finding⁵ based on studies of ancient metal grains. Our result, however, is more difficult to understand within the context of current models, because there is no clear way to make ^{16}O more abundant in Solar System rocks than in the Sun.

The lunar regolith is exposed to solar wind and ‘gardened’ by meteorite impacts such that particles are continually cycled through the surface; in some soils, nearly all grains have been exposed to solar wind, and some grains have probably seen several exposures⁶. The use of lunar soil for understanding solar wind compositions of various elements is well established, but the deconvolution of contributions from various sources (for example, solar wind, meteorite, cometary, ‘planetary’ and cosmogenic) has proven to be difficult. One distinguishing characteristic of solar wind is that the energetic atoms are implanted below the mineral surface. The solar wind also sputters (erodes) surfaces, resulting in ejection of surface atoms, causing damage to the surface layers, and potentially mixing components.

Isotopic measurements reveal a complex interplay between many components and the nature of the target minerals. For instance, Hashizume and colleagues have shown that a pyroxene grain in lunar regolith breccia 79035 has low ^{13}C (by up to 10%; ref. 7), and low ^{15}N (by 30%; ref. 8). These are correlated with low D/H and ascribed to the solar wind. In contrast, an ilmenite from another lunar breccia, 71501, shows depletion in ^{13}C , but not the low D/H or ^{15}N , and this may be due to a silicon-bearing coating preserving non-solar components.

In order to measure oxygen isotopes, a target with low intrinsic oxygen is required. We recovered six metal spherules (‘grains’) from a recent lunar soil, 10084,85 (that is, sample 10084, split 85). The grains range in size from 20 to 80 μm , and we mounted five of these grains (one was lost in handling) in gold foil along with polished lunar ilmenite grains. We analysed three of the metal grains: two grains were not analysed—one was intentionally preserved, and the other was unsuitable for analysis. The data table of all analyses is available as Supplementary Information.

Our prediction, before analysis⁹, was that a high initial O count rate would reflect surface oxidation (terrestrial alteration), followed by a rapid decrease in the count rate where the subsurface solar wind

should reside. Two iron metal grains, denoted α and β , show the expected steep O concentration gradients over the first 20 nm and then a more gradual decay (Fig. 1). High-precision measurements show that for one of the grains, $\Delta^{17}\text{O}$ is constant and clearly positive. Oxygen three-isotope plots are presented in Fig. 2 for all grains analysed in this study, combining three different methods of analysis. A single-collector measurement starts on the terrestrial mass-dependent (slope 0.5) fractionation line (TFL) and becomes progressively heavier in both $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$. This trend continues with subsequent multiple-ion-counter measurements. Finally, the highest-precision analyses using a Faraday cup and two ion counters show clear resolution of the implanted oxygen isotopic composition from the TFL. Nine high-precision analyses of grain α lie on a mass-dependent-fractionation line (αFL) that is offset from the TFL with $\Delta^{17}\text{O}$ of $+25.6 \pm 3.2\text{‰}$ ($2\sigma_m$; MSWD = 1.1), the error in the mean being 16σ removed from the TFL (here σ_m is the standard error of the mean, MSWD is mean squared weighted deviates). In comparison, the lunar ilmenite $\Delta^{17}\text{O}$ values, representing indigenous lunar oxygen, are all within 3σ of the TFL.

Most data from grain α lie close to the intercept of the αFL with a line of slope 1 corresponding to variation in ^{16}O (^{16}OFL ; Fig. 2). This suggests that the subsurface (solar-wind) component is represented by a composition near $\delta^{17}\text{O} = \delta^{18}\text{O} = +50\text{‰}$. The data lying significantly to the right of this point are from a single set of analyses that correspond to depths of 180–360 nm and may represent mass-dependent fractionation associated with solar energetic particles.

We believe, on the basis of the following arguments, that the exotic oxygen revealed in these grains is that of the solar wind. Lunar soil 10084 has one of the highest solar wind exposures known, on the basis of its high ^{36}Ar concentration ($2.1 \times 10^{-7} \text{ cm}^3 \text{ STP cm}^{-2}$) and percentage of plagioclase grains being track-rich (96%)¹⁰. With such a high fraction of irradiated grains, it would be highly unlikely that any grain has escaped exposure to the solar wind. If the 10084 metal grains have been exposed to solar wind, the exotic oxygen must be solar because no other oxygen component is present. Furthermore, a relatively high concentration of ^4He in ilmenite and high $^4\text{He}/^{20}\text{Ne}$ are consistent with extensive recent exposure¹¹.

Oxygen is implanted to similar depths and concentrations as solar wind H, C and N isotopes in silicate grains^{8,12}, but it is apparent that all elements show deeper penetration than expected from solar wind implantation models¹³. At this stage, it is not clear whether such discrepancies are significant. For our measurements, the grains have been distorted from pressing into gold. Imperfect focusing of the primary beam can also cause apparent broadening of profiles, but it is also possible that thermal relaxation (volume diffusion) has affected the distributions. Hydrogen isotopes in these metal grains are normal, as are those in 79035 (ref. 7), due to rapid diffusion of terrestrial H through the metal. The diffusion rate of oxygen through iron is difficult to determine in part due to the formation of an oxide

¹Research School of Earth Sciences, The Australian National University, Canberra, Australian Capital Territory 0200, Australia. ²Lunar and Planetary Institute, 3600 Bay Area Boulevard, Houston, Texas 77058, USA. [†]Present address: School of Geography and Environmental Science, Monash University, Melbourne, Victoria 3800, Australia.

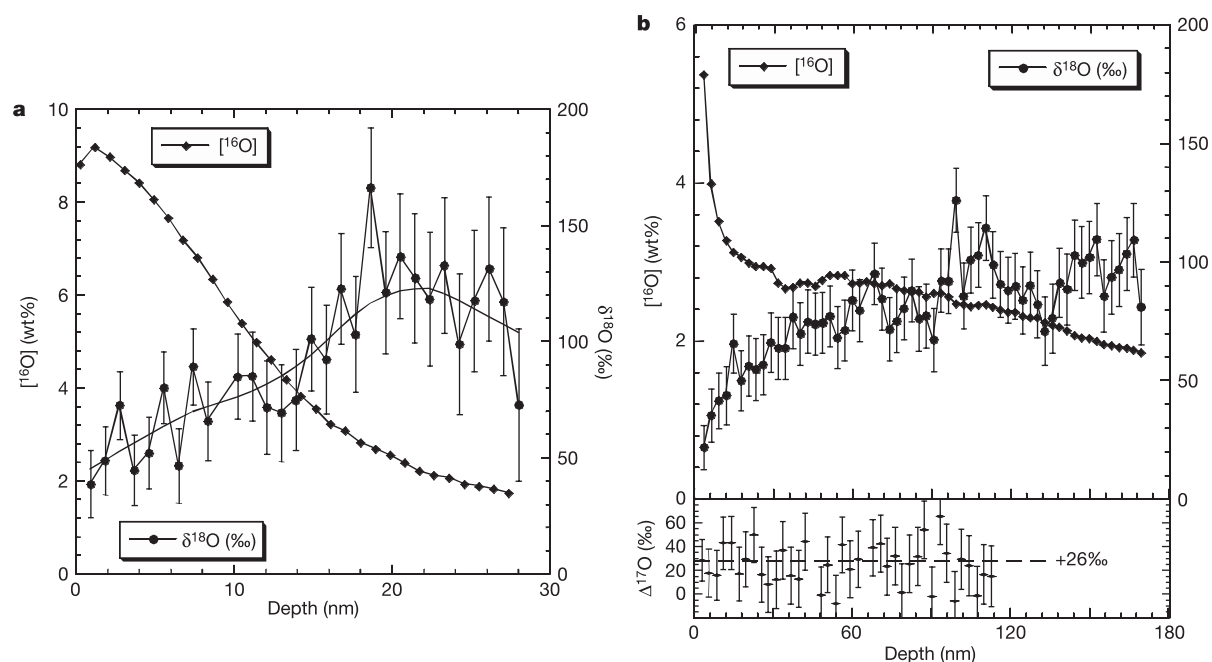


Figure 1 | Oxygen-isotope depth profiles for two iron grains from lunar soil 10084. **a**, A high-resolution profile for grain β in single-ion-counter collection mode (primary Cs^+ beam 0.6 nA) shows a surface layer less than 15 nm thick with a maximum oxygen concentration of ~ 9 wt%. The oxygen signal quickly decays to a level around 1–2 wt%; $\delta^{18}\text{O}$ systematically rises during the analysis. **b**, Multiple collection measurements (Faraday cup and two ion counters; 2 nA Cs^+ primary ion beam) of grain α also shows a high surface oxygen concentration that is again rapidly sputtered away with a corresponding increase in $\delta^{18}\text{O}$. For this analysis, the $\Delta^{17}\text{O}$ determination is sufficiently precise to show constancy with depth around the mean $\Delta^{17}\text{O}$ obtained for all analyses of grain α (dashed line at +26‰, see Fig. 2). The concentration profiles are broader than expected for solar wind

layer that inhibits diffusion¹⁴. Nevertheless, some experiments indicate significant diffusion of oxygen through iron¹⁵ and this is compatible with the perturbation of the depth profiles. Diffusion could result in higher oxygen concentrations at the surface, and this is supported by the constancy of oxygen isotopic composition during the depth profile shown in Fig. 1.

The oxygen profiles we observe from a contemporary lunar soil (10084) are very different to those measured from an ancient lunar regolith (79035; ref. 5). The compaction age of 1–2 Gyr (refs 12, 16) represents the timing of cessation of solar wind exposure and the presence of thick oxide layers on iron grains from 79035 indicates an extended thermal history in an oxidizing environment subsequent to compaction. It should be expected that these grains have been compromised in terms of their solar wind retention; Hashizume and Chaussidon⁵ identify the source of anomalous oxygen they measure as solar energetic particles (SEP) because of the large measurement depths (up to 1 μm) below the grain surfaces. In most of the grains they analysed, oxygen lies close (within 2σ error limits) to the TFL for the duration of analysis⁵. In only a few of the grains, the composition migrates from the TFL to both positive and negative $\Delta^{17}\text{O}$ values. A fractionated composition with $\Delta^{17}\text{O}$ of -20‰ is favoured as the SEP contribution because the concentration is the lowest observed (<1.4 wt%). Extrapolation of this fractionated composition back to a slope-1 ^{16}O FL yields the inferred composition of the Sun at $-40 \pm 8\text{‰}$, or lower.

An accurate assessment of the 79035 oxygen isotope data is particularly difficult because of the deconvolution of multiple components present in that sample. For two of the iron grains analysed from 10084, a constant and positive $\Delta^{17}\text{O}$ is measured and this is the only composition measured at the depths expected for

implantation alone. Using the SOHO solar-wind-velocity database²⁵ and the TRIM ion implantation programme¹³, we find that for the average velocity of solar wind (430 km s^{-1} , 16 keV O ions), a peak implantation depth for oxygen in iron metal is only 17 nm and even for the extreme velocities ($1,000 \text{ km s}^{-1} = 80 \text{ keV O ions}$), the peak implantation depth is only 90 nm. This broadening may be related to oxygen diffusion following solar wind implantation. Measured oxygen isotope compositions are expressed in standard delta notation relative to standard mean ocean water (SMOW), such that $\delta^i\text{O} = ({}^i\text{O}/{}^{16}\text{O})_{\text{meas}}/({}^i\text{O}/{}^{16}\text{O}_{\text{SMOW}}) - 1 \times 1,000$, for $i = 17$ and 18. Deviation of compositions from the terrestrial mass fractionation line are expressed in terms of the effective residual ^{17}O expressed as $\Delta^{17}\text{O} = \delta^{17}\text{O} - 0.52\delta^{18}\text{O}$. Error bars are 1σ .

solar wind implantation. The extreme data from 79035 are inferred to represent SEP and require extrapolation back to a model dependent composition; the data from 10084 indicate a specific composition for the solar wind and hence the Sun.

The result presented here is controversial because it does not place the solar oxygen isotope composition within the previously known range of meteoritic ^{16}O variations—that is, close to that of refractory inclusions with ^{16}O excess of +50‰, or close to that of the planetary bodies with near-terrestrial compositions. Our measurements extend the ^{16}O -fractionation range observed in the Solar System, with the Sun apparently enriched in ^{17}O and ^{18}O by around 50‰. As such, planets and meteorite refractory inclusions must represent highly fractionated compositions relative to the starting protosolar nebula.

Our estimate for the solar oxygen isotope composition lies on the ^{16}O FL, which is consistent with the variations observed in the Solar System. This variability was originally attributed to admixture of varying amounts of an ^{16}O -rich nucleosynthetic component¹⁷. A chemical mechanism has also been proposed, although no specific mechanism to isolate the fractionated components has been substantiated^{18–20}. Recent attention has focused on models based around ultraviolet dissociation of CO with the fractionation caused by self-shielding of abundant C^{16}O (and continued dissociation of C^{17}O and C^{18}O at greater optical depth). Dissociated oxygen quickly reacts with H on mineral surfaces, so the gas (CO) becomes enriched in C^{16}O and solid phases (ice) in ^{17}O and ^{18}O . In current versions of this model^{21–23}, the Sun should be ^{16}O -rich, close to the compositions of refractory inclusions. If our measurements accurately reflect the composition of the Sun (^{17}O - and ^{18}O -rich), these models in their current form cannot be correct. Alternatively, the solar wind may not represent the bulk composition of the Sun, but rather reflect later

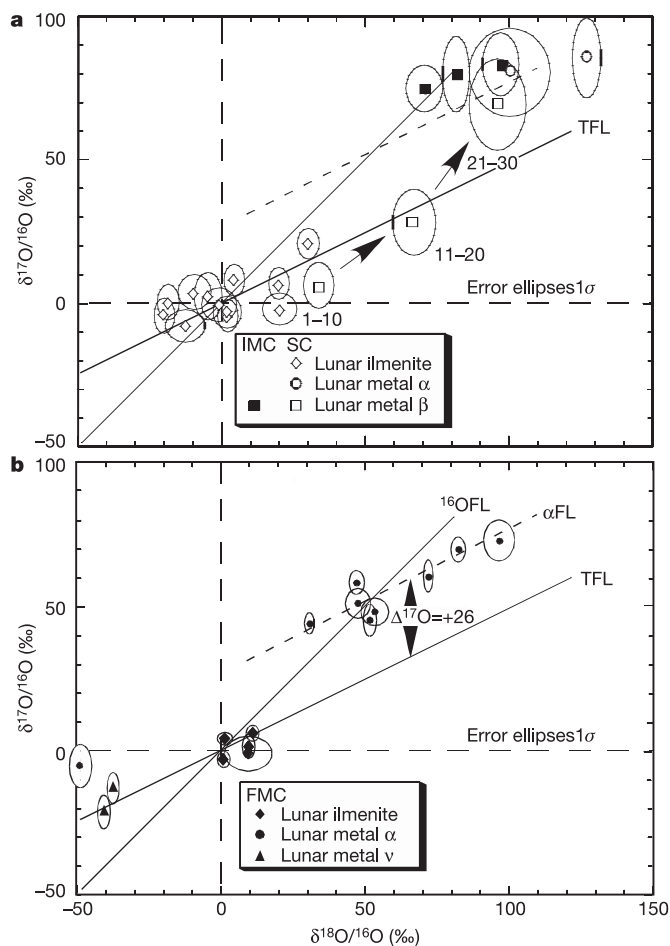


Figure 2 | Three-isotope plots of oxygen isotope compositions in lunar metal grains and lunar ilmenites. **a**, The single-collector (SC) high-resolution profile of grain β shows a systematic change in isotopic composition first along and then away from the slope-0.5 terrestrial mass-dependent fractionation line (TFL) towards elevated $\delta^{18}\text{O}$, $\delta^{17}\text{O}$ and $\Delta^{17}\text{O}$. SC data are grouped in subsets of 10 from the profile of Fig. 1 to yield higher precision in the $\delta^{17}\text{O}$ measurement. Subsequent multiple collection measurements (three ion counters, IMC) yield higher precision and continue the trend to elevated $\delta^{18}\text{O}$ and $\delta^{17}\text{O}$, showing clear resolution from the TFL. **b**, Multiple collection measurements with Faraday cup and two ion counters (FMC) yield the highest-precision data, and oxygen isotope data from grain α is seen to lie on a single mass-dependent fractionation line (αFL) offset from the TFL by $+25.6 \pm 3.2\text{‰}$ ($2\sigma_m$; MSWD = 1.1). The concentration of data around the intersection of the αFL and the slope-1 ^{16}OFL (ref. 26) suggests the main O isotopic variation is in the ^{16}O abundance. In this case the calculated ^{16}O excess²⁷ is $-53.4 \pm 4.7\text{‰}$ ($2\sigma_m$; MSWD = 2.3). Being the only isotopic component beneath the surface of these grains, this O isotopic composition must be that of the solar wind and hence the Sun. Also shown are data from lunar ilmenites that represent indigenous lunar oxygen and reproducibility of analytical protocols, and analyses of grain ν and one of α that had high O count rates related to surface oxide layers. Error ellipses are 1σ .

infall, for example of nebula water, which is expected to be enriched in ^{17}O and ^{18}O (refs 4, 24). Some of these issues may be resolved with the analysis of samples returned from the Genesis mission.

Received 3 August 2005; accepted 26 January 2006.

1. Clayton, R. N. Oxygen isotopes in meteorites. *Annu. Rev. Earth Planet. Sci.* **21**, 115–149 (1993).

2. Birck, J. L. An overview of isotopic anomalies in extraterrestrial materials and their nucleosynthetic heritage. *Rev. Mineral. Geochem.* **55**, 25–64 (2004).
3. Wiens, R. C., Bochsler, P., Burnett, D. S. & Wimmer-Schweingruber, R. F. Solar and solar wind isotopic compositions. *Earth Planet. Sci. Lett.* **226**, 549–565 (2004).
4. Wiens, R. C., Huss, G. R. & Burnett, D. S. The solar oxygen-isotopic composition: Predictions and implications for solar nebula processes. *Meteorit. Planet. Sci.* **34**, 99–107 (1999).
5. Hashizume, K. & Chaussidon, M. A non-terrestrial ^{16}O -rich isotopic composition for the protosolar nebula. *Nature* **434**, 619–622 (2005).
6. Pillinger, C. T. Solar-wind exposure effects in lunar soil. *Rep. Prog. Phys.* **42**, 897–968 (1979).
7. Hashizume, K., Chaussidon, M., Marty, B. & Terada, K. Protosolar carbon isotopic composition: Implications for the origin of meteoritic organics. *Astrophys. J.* **600**, 480–484 (2004).
8. Hashizume, K., Marty, B. & Wieler, R. Analyses of nitrogen and argon in single lunar grains: towards a quantification of the asteroidal contribution to planetary surfaces. *Earth Planet. Sci. Lett.* **202**, 201–216 (2002).
9. Ireland, T. R., Holden, P., Norman, M. D. & Clarke, J. Oxygen isotopes in lunar metal grains: A natural Genesis experiment. *Lunar Planet. Sci. XXXV*, abstr. 1448 (2004).
10. Wieler, R., Etique, P., Signer, P. & Poupeau, G. In *Proc. 11th Lunar and Planetary Science Conference* (ed. Merrill, R. B.) 1369–1393 (Pergamon, New York, 1980).
11. Eugster, O., Grögler, N., Mendia, M. D., Eberhardt, P. & Geiss, J. Trapped solar wind noble gases and exposure age of Luna 16 lunar fines. *Geochim. Cosmochim. Acta* **37**, 1991–2003 (1973).
12. Hashizume, K., Chaussidon, M., Marty, B. & Robert, F. Solar wind record on the Moon: Deciphering presolar from planetary nitrogen. *Science* **290**, 1142–1145 (2000).
13. Ziegler, J. F., Biersack, J. P. & Littmark, U. *The Stopping and Range of Ions in Solids* (Pergamon, New York, 1985).
14. Martin, M., Mader, W. & Fromm, E. Oxidation of iron, aluminium and titanium films in the temperature range 50–200 °C. *Thin Solid Films* **250**, 61–66 (1994).
15. Takada, J., Yamamoto, Y. & Adachi, M. Diffusion coefficient of oxygen in alpha-iron determined by internal oxidation technique. *Z. Metallk.* **77**, 6–11 (1986).
16. Wieler, R., Kehm, K., Meshik, A. P. & Hohenberg, C. M. Secular changes in the xenon and krypton abundances in the solar wind recorded in single lunar grains. *Nature* **384**, 46–49 (1996).
17. Clayton, R. N., Grossman, L. & Mayeda, T. K. A component of primitive nuclear composition in carbonaceous meteorites. *Science* **182**, 485–488 (1973).
18. Nuth, J. A. I., Hallenbeck, S. L. & Rietmeijer, F. J. M. Interstellar and interplanetary grains: Recent developments and new opportunities for experimental chemistry. *Earth Moon Planets* **80**, 73–112 (1998).
19. Thieme, M. H. Mass-independent isotope effects in planetary atmospheres and the early solar system. *Science* **283**, 341–345 (1999).
20. Thieme, M. H. & Heidenreich, J. E. I. The mass-independent fractionation of oxygen—a novel isotope effect and its possible cosmochemical implications. *Science* **219**, 1073–1075 (1983).
21. Clayton, R. N. Self-shielding in the solar nebula. *Nature* **415**, 860–861 (2002).
22. Yurimoto, H. & Kuramoto, K. Molecular cloud origin for the oxygen isotope heterogeneity in the solar system. *Science* **305**, 1763–1766 (2004).
23. Lyons, J. R. & Young, E. D. CO self-shielding as the origin of oxygen isotope anomalies in the early solar nebula. *Nature* **435**, 317–320 (2005).
24. Young, E. D. in *Workshop on Oxygen in the Early Solar System* Abstr. 9010 (Gatlinburg, Tennessee, 2005); available at (<http://www.lpi.usra.edu/meetings/ess2005/>) (2005).
25. CELIAS/MTOF experiment on the Solar Heliospheric Observatory (SOHO). (<http://umtof.umd.edu/pub/>) (2005).
26. Young, E. D. & Russell, S. S. Oxygen reservoirs in the early solar nebula inferred from an Allende CAI. *Science* **282**, 452–455 (1998).
27. Clayton, R. N. & Mayeda, T. K. The oxygen isotope record in Murchison and other carbonaceous chondrites. *Earth Planet. Sci. Lett.* **67**, 151–161 (1984).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements T. Senden carried out the atomic force microscopy of ion probe pits. NASA provided lunar soil 10084.85. This work was supported by the Australian Research Council. We thank M. Thieme, K. McKeegan and U. Wiechert for discussions, and K. Hashizume, G. Huss and A. Davis for comments and suggestions.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.R.I. (Trevor.Ireland@anu.edu.au).

Quantum interference between two single photons emitted by independently trapped atoms

J. Beugnon¹, M. P. A. Jones¹, J. Dingjan¹, B. Darquié¹, G. Messin¹, A. Browaeys¹ & P. Grangier¹

When two indistinguishable single photons are fed into the two input ports of a beam splitter, the photons will coalesce and leave together from the same output port. This is a quantum interference effect, which occurs because two possible paths—in which the photons leave by different output ports—interfere destructively. This effect was first observed in parametric downconversion¹ (in which a nonlinear crystal splits a single photon into two photons of lower energy), then from two separate downconversion crystals², as well as with single photons produced one after the other by the same quantum emitter^{3–6}. With the recent developments in quantum information research, much attention has been devoted to this interference effect as a resource for quantum data processing using linear optics techniques^{2,7–11}. To ensure the scalability of schemes based on these ideas, it is crucial that indistinguishable photons are emitted by a collection of synchronized, but otherwise independent sources. Here we demonstrate the quantum interference of two single photons emitted by two independently trapped single atoms, bridging the gap towards the simultaneous emission of many indistinguishable single photons by different emitters. Our data analysis shows that the observed coalescence is mainly limited by wavefront matching of the light emitted by the two atoms, and to a lesser extent by the motion of each atom in its own trap.

A basic requirement for most quantum computing schemes is the implementation of two-qubit quantum gates¹². If the qubits are encoded in single photons, the gate can be obtained by using an interference effect between the photons, followed by a measurement-induced state projection⁹. One may also use qubits encoded in solid-state systems such as quantum dots¹³, or in atomic systems such as ions¹⁴ or neutral atoms¹⁵. One way to entangle the atomic qubits without direct interaction, and thus realise quantum gates, is to use them as single photon sources, so that the emitted photons are entangled with the internal states of the emitters. The interference of two photons emitted by such sources projects the state of the two atoms into an entangled state¹⁶. Many protocols based on this conditional entanglement have been proposed^{10,11}, and experimental work is under way to implement them¹⁷. The photons involved in such schemes do not need to be initially entangled, and can even be emitted by different sources², but they need to be indistinguishable. However, it is in general not easy to have several (possibly many) independent sources emitting indistinguishable photons. With quantum dots in microcavities^{3,6}, the dispersion in frequency associated with differences in fabrication is usually much too large for the photons to be emitted at the same frequency. With atoms in cavities⁴, each emitter is itself a complicated experiment, and cannot easily be multiplied. In this paper we address this problem by using two single atoms in two neighbouring traps emitting in free space, and we demonstrate that these atoms do emit indistinguishable photons. This scheme can easily be scaled to arrays of traps¹⁸.

Our experiment uses two single rubidium-87 atoms, confined in separate optical dipole traps. These traps are formed in the focal plane of the same high-numerical aperture lens, and loaded from a cloud of cold atoms in an optical molasses¹⁹. The two traps, each of which has a waist of 1 μm , are separated by a distance of 6 μm . To obtain triggered single photon emission from the two atoms, we simultaneously excite them with a high efficiency ($>95\%$) using a σ^+ -polarized pulsed laser beam which drives the $F = 2$, $m_F = 2$ to $F' = 3$, $m_F' = 3$ closed transition²⁰. The quantization axis is defined by a magnetic field of several gauss. Both atoms are excited by the

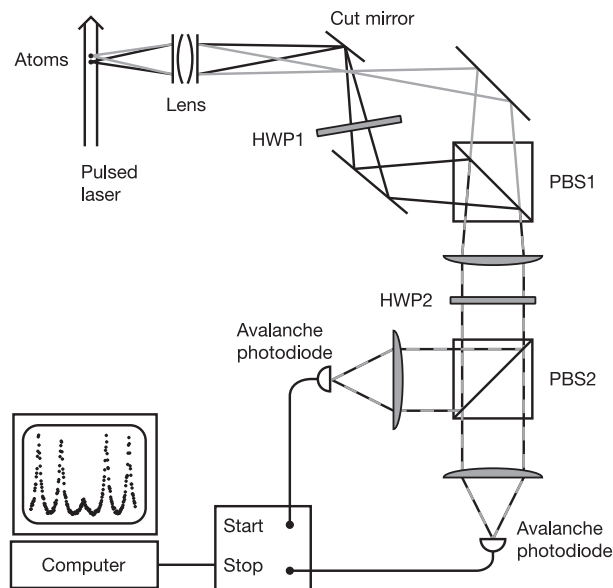


Figure 1 | Experimental set-up. The two atoms are trapped in two dipole traps separated by 6 μm , and they are excited by the same pulsed laser beam. The trap depth is 1.5 mK, and the trap frequency along the axis of the pulsed laser beam is 120 kHz. The emitted photons are collected by the same lens that is used to create the dipole traps. The light from one of the traps is separated off using a cut mirror placed close to the image plane of the objective. In the plane of the cut mirror, the spot corresponding to each trap has a waist of $\sim 90 \mu\text{m}$, and the two images are separated by 500 μm . The half-wave plate HWP1 is oriented such that, at polarizing beam splitter PBS1, the light beams from the two atoms are recombined into the same spatial mode, but with orthogonal polarizations. There are then two configurations to detect the photons: either the axis of HWP2 is set so that the two orthogonal incident polarizations are equally mixed in each output of PBS2, as in a 50/50 beam splitter, or the axis is set so that the polarizations are unchanged, and then PBS2 simply separates the two beams coming from the two atoms without mixing them. Two avalanche photodiodes are placed in the two output ports of PBS2. The measured overall collection and detection efficiency is 0.6% for each photodiode²¹.

¹Laboratoire Charles Fabry de l'Institut d'Optique (UMR 8501), Bâtiment 503, Centre Universitaire, 91403 Orsay cedex, France.

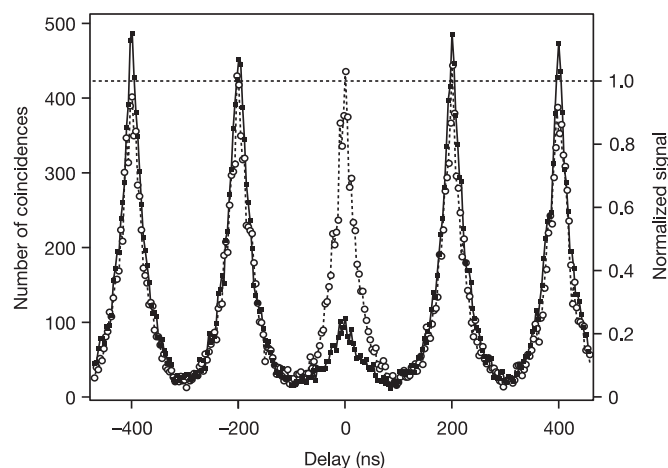


Figure 2 | Histograms of the time delays of the arrival of two photons on the avalanche photodiodes, in the start-stop configuration. Black squares correspond to the 50/50 beam-splitter configuration (interfering beams). Open circles correspond to the beam-separator configuration (independent beams). These histograms have been binned three times, leading to a 3.6-ns resolution. The total accumulation time is about five hours, corresponding to about 6,600 events with two photons arriving on the beam splitter at around zero delay. The solid and dashed lines are a guide to the eye. The normalized signal is obtained by dividing the number of counts by the average value of the peak height in the beam-separator configuration.

same short (less than 4 ns) π -pulse. Each one then spontaneously emits a single photon²¹ with a lifetime of 26 ns. The photons are collected by the same objective lens that is used to focus the dipole trap beams¹⁹, and detected using a pair of avalanche photodiodes. Between the objective and the photodiodes, an optical set-up composed of two half-wave plates and two polarizing beam-splitter cubes (HWP1, HWP2 and PBS1, PBS2) is inserted in the beam path (see Fig. 1). It can be configured either as a 50/50 beam splitter that mixes the light from the two atoms on each detector, or as a beam separator that sends the light from each atom to only one of the detectors. The avalanche photodiodes are connected to a high-resolution counting card in a start-stop configuration. This allows us to measure the number of coincident photodetections as a function of the delay between the arrival times of the two photons on the photodiodes, with a resolution of about 1.2 ns. In the 50/50 beam-splitter configuration, the detectors cannot distinguish which atom has emitted a photon, and we expect to observe the coalescence effect. In the beam-separator configuration, each avalanche photodiode monitors only the light emitted by one of the two atoms, and coincidence counts can be due only to independent emissions by both atoms.

The measurements are performed by repeating the following procedure. First, we detect the simultaneous presence of one atom in each trap in real time by measuring their fluorescence from the molasses light used to load the traps. We then trigger a sequence that alternates a burst of 575 pulsed excitations, lasting 115 μ s, with a 885- μ s cooling period using the molasses light. This alternation is repeated 15 times before stopping and recapturing a new pair of atoms. During the excitation periods, the π -pulses irradiate both atoms every 200 ns, and the counting card accumulates the number of double detections produced by the two avalanche photodiodes. This sequence maximizes the number of single photons that we can obtain before the two atoms are heated out of the trap²¹. After the 15 bursts of pulsed excitations, we measured a probability of 65% to keep each atom in its trap. At the end of the sequence, we switch the molasses back on and wait on average about 300 ms until we detect two atoms again. Two histograms are accumulated for the same

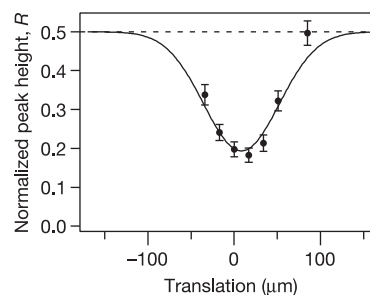


Figure 3 | Influence of wavefront matching. The ratio of the height of the residual peak at zero delay in the beam-splitter configuration to the average height of the peaks in the beam-separator configuration, as a function of the relative distance between the two beams, translated parallel to each other. The solid curve is the expected ratio, calculated from the measured beam waist of the two beams. The amplitude and the centre of this curve are adjusted to fit the data. Error bars show ± 1 s.d.

number of repetitions: one in the 50/50 beam-splitter configuration, and one in the beam-separator configuration.

The two histograms are shown in Fig. 2, without background subtraction. Both histograms consist of a series of peaks separated by 200 ns, the repetition period of the pulsed laser. The width of the peaks is determined by the 26-ns lifetime of the excited state. In the beam-separator configuration, all peaks are identical, and always correspond to a double detection with one photon coming from each atom. Hence, their height gives a natural calibration of the experiment. A histogram in the 50/50 beam-splitter configuration can be normalized by dividing by the height of the peaks in its corresponding histogram measured in the beam-separator configuration. The normalized signals that are then independent of collection efficiency, detection efficiency and experiment duration, and allow histograms taken under different conditions to be compared. In the 50/50 beam-splitter configuration, the peak at zero delay is clearly much smaller than the other peaks. As each atom is a very good source of single photons²¹, this peak also consists only of events where both atoms have emitted a photon. In contrast, the other peaks consist of events where two photons are successively emitted, either by the same atom, or by both atoms. Because the peaks at non-zero delay are almost the same in both configurations, we can deduce that almost all registered counts are due to events where both atoms were present.

In the case of perfect coalescence, the peak at zero delay in the 50/50 beam-splitter configuration would be absent: as the two photons leave via the same port, there can be no coincidences. We attribute the residual peak that we observe in Fig. 2 to an imperfect overlap of the spatial modes of the two photons, which then do not interfere. To illustrate this effect experimentally, we varied the overlap between the two modes in a controlled way by translating one beam relative to the other (translation of the cut mirror; see Fig. 1). Figure 3 shows the normalized height R of the residual peak at zero delay, as a function of the separation between the two images. For a given spatial overlap K between the amplitude of the two modes, the ratio R is expected to be $(1-K^2)/2$ (see Methods section and Supplementary Information). The solid line is a gaussian fit based on the experimental value of the beam size in the image plane, and considering the maximal wavefront overlap $K_{\max}$ as an adjustable parameter. The agreement with the coincidence data is very good, which confirms the crucial role of good mode matching of the two beams in our experiment. We obtain from the fit the maximum wavefront overlap $K_{\max} = 0.78 \pm 0.03$. This imperfect overlap is consistent with the errors we measure on the beam positions and waists.

Finally, we analyse the structure of the time spectrum around zero delay. The small peak at zero delay from Fig. 2 is displayed on a larger scale in Fig. 4. The dashed line corresponds to a model where the

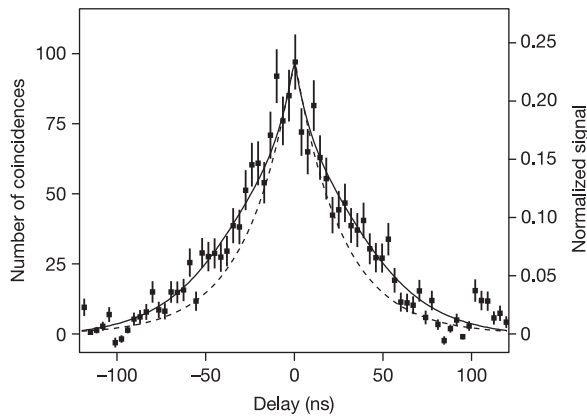


Figure 4 | Zoom of the histogram of Fig. 2 in the 50/50 beam-splitter configuration, around zero delay. This curve is obtained from Fig. 2 by subtraction of the contribution from the background and neighbouring peaks. The squares represent the experimental data. The solid line is a fit by the model described in the Methods section, taking into account the finite temperature of the atoms and the spatial overlap. The dotted line is the expected signal for zero temperature. The error bars correspond to $\pm\sqrt{N}$ statistical photon counting noise.

wavepackets of the two photons are identical and arrive at the same time on the beamsplitter, but with imperfect spatial overlap of the two beams. This curve does not correctly reproduce the experimental data: the experimental dots seem to sit on a slightly wider curve. Due to their finite temperature, the atoms move in the trapping potential and experience a range of light-shifts. This changes their internal energy, and thus modifies the frequency of the emitted photon. For two photons at different frequencies, the correlation signal would exhibit a beat note as already seen in ref. 4. If the two photons now have a distribution of frequencies, the correlation signal consists of a beat note averaged out over this distribution. This gives rise to a slightly broader structure for the signal, which is well fitted by the solid line predicted by our simple model (see details in the Methods section).

By fitting the experimental data shown in Fig. 4 with our model, we extract the overlap of the spatial modes $K = 0.7 \pm 0.05$ and the temperature of the atoms $T = 180 \pm 20 \mu\text{K}$. In a separate experiment, we measured the temperature of the atoms in the dipole trap, which is initially close to $120 \pm 10 \mu\text{K}$. Each pulse followed by the spontaneous emission of the photon increases the energy of the atom by one recoil. We calculate that after the first $115 \mu\text{s}$ of pulsed excitations the temperature rises by $60 \mu\text{K}$, in good agreement with the temperature obtained from the fit above. We also checked experimentally that each cooling period resets the temperature of the atom to its initial value. A comparison of the fit with the dashed curve, which corresponds to zero temperature, confirms that at present the imperfect interference is due mainly to the imperfect optical wavefront matching, and not to the motion of the atoms in the traps.

Thus we have experimentally demonstrated the coalescence of two photons emitted by two independent trapped atoms. The contrast of the interference is limited by the overlap (in free space) of the spatial modes of the fluorescence light emitted by the two atoms. By coupling the light from each of the atoms into identical single-mode optical fibres, this overlap could be greatly improved, though this may be at the cost of a reduced counting rate. The shape of the residual signal around zero delay is well explained by a broadening due to the finite temperature of the atoms in the trap. Better wavefront overlap and further cooling of the atoms will improve the overall quality of this interference and will make this system suitable as a resource for entangling two atoms.

METHODS

Derivation of the experimental signal. If $f_k(\mathbf{r})\varepsilon_k(t)$ is the field emitted by the atom k ($k = 1, 2$), ref. 22 shows that the probability of detecting one photon at $\mathbf{r}_A$ and of detecting the other one at $\mathbf{r}_B$, in the 50/50 beam-splitter configuration, after a delay τ , is proportional to:

$$w^{(2)}(\tau, \mathbf{r}_A, \mathbf{r}_B) = \int |f_1(\mathbf{r}_B)f_2(\mathbf{r}_A)\varepsilon_1(t+\tau)\varepsilon_2(t) - f_2(\mathbf{r}_B)f_1(\mathbf{r}_A)\varepsilon_2(t+\tau)\varepsilon_1(t)|^2 dt$$

which can be understood as the interference of two paths. Assuming a temporal form of the field $\varepsilon_k(t) = H(t)e^{-\frac{t}{\tau}}e^{i\omega_k t}$ where $H(t)$ is the step function, we obtain:

$$w^2(\tau) \propto e^{-\Gamma|\tau|}(1 - K^2 \cos \Delta\omega\tau)$$

where $K = |\int d\mathbf{r} f_1^*(\mathbf{r})f_2(\mathbf{r})| / \sqrt{\int d\mathbf{r} |f_1(\mathbf{r})|^2 \times \int d\mathbf{r} |f_2(\mathbf{r})|^2}$ is the spatial overlap of the electric field, and $\Delta\omega$ is the frequency difference between the two emitted photons. The double detection probability for $\tau = 0$ is proportional to $(1-K^2)$, and so is the normalized ratio R defined in the text. The proportionality factor is determined in the absence of interferences ($K = 0$): in this case, the two photons behave as distinguishable particles and have a probability of 1/2 of leaving the 50/50 beam splitter through two different ports. Thus, the normalized ratio R is $(1-K^2)/2$.

Model including the finite temperature of the atoms in the trap. To take into account the finite temperature of the atoms in the trap, we integrate the expected signal for two photons interfering with a frequency difference $\Delta\omega$ and a spatial overlap K (see above), over the probability distribution of frequency differences. To obtain this probability distribution, we solve the equations of motion for a thermal ensemble of single atoms in the trap experiencing pulsed excitations during $115 \mu\text{s}$, followed by a decay in a random direction. After each pulse, we calculate the lightshifts of all the atoms in the ensemble. We repeat this for 575 pulses to obtain the distribution of lightshifts. This distribution is found to be well represented by a function of the form $U^2 e^{-\frac{\Delta\omega}{U}}$. The value of $\Delta\omega$ is proportional to the difference in lightshifts experienced by the atoms when they emit the photons. We then calculate the auto-correlation of the lightshift distribution to get the probability distribution of $\Delta\omega$. By averaging over this distribution of lightshift differences, we obtain the normalized coincidence rate signal as an analytical function with only two fitting parameters, the spatial overlap and the temperature of the atoms.

Received 25 October 2005; accepted 31 January 2006.

- Hong, C. K., Ou, Z. Y. & Mandel, L. Measurement of subpicosecond time intervals between two photons by interference. *Phys. Rev. Lett.* **59**, 2044–2046 (1987).
- de Riedmatten, H., Marcikic, J., Tittel, W., Zbinden, M. & Gisin, N. Quantum interference with photon pairs created in spatially separated sources. *Phys. Rev. A* **67**, 022301 (2003).
- Varoutsis, S. *et al.* Restoration of photon indistinguishability in the emission of a semiconductor quantum dot. *Phys. Rev. B* **72**, 041303(R) (2005).
- Legero, T., Wilk, T., Hennrich, M., Rempe, G. & Kuhn, A. Quantum beat of two single photons. *Phys. Rev. Lett.* **93**, 070503 (2004).
- Kiraz, A. *et al.* Indistinguishable photons from a single molecule. *Phys. Rev. Lett.* **94**, 223602 (2005).
- Santori, C., Fattal, D., Vučković, J., Salomon, G. S. & Yamamoto, Y. Indistinguishable photons from a single-photon device. *Nature* **419**, 594–597 (2002).
- Dowling, J. P., Franson, J. D., Lee, H. & Milburn, G. J. Towards scalable linear-optical quantum computers. *Quantum Inf. Process.* **3**, 205–213 (2004).
- Pan, J.-W., Bouwmeester, D., Weinfurter, H. & Zeilinger, A. Experimental entanglement swapping: entangling photons that never interacted. *Phys. Rev. Lett.* **80**, 3891–3894 (1998).
- Knill, E., Laflamme, R. & Milburn, G. J. A scheme for efficient quantum computation with linear optics. *Nature* **409**, 46–52 (2001).
- Lim, Y. L., Beige, A. & Kwek, L. C. Repeat-until-success linear optics distributed quantum computing. *Phys. Rev. Lett.* **95**, 030505 (2005).
- Barrett, S. D. & Kok, P. Efficient high-fidelity quantum computation using matter qubits and linear optics. *Phys. Rev. A* **71**, 060310 (2003).
- Zoller, P., Cirac, J. I., Duan, L. & García-Ripoll, J. J. in *Les Houches, Session LXXIX, 2003—Quantum Entanglement and Information Processing* (eds Estève, D., Raimond, J.-M. & Dalibard, J.) 187–222 (Elsevier, Amsterdam, 2004).
- Paillard, M. *et al.* Spin relaxation quenching in semiconductor quantum dots. *Phys. Rev. Lett.* **86**, 1634–1637 (2001).
- Monroe, C., Meekhof, D. M., King, B. E., Itano, W. M. & Wineland, D. J. Demonstration of a fundamental quantum logic gate. *Phys. Rev. Lett.* **75**, 4714–4717 (1995).
- Schrader, D. *et al.* Neutral atom quantum register. *Phys. Rev. Lett.* **93**, 150501 (2004).

16. Simon, C. & Irvine, W. T. M. Robust long-distance entanglement and a loophole-free Bell test with ions and photons. *Phys. Rev. Lett.* **91**, 110405 (2003).
17. Blinov, B. B., Moehring, D. L., Duan, L.-M. & Monroe, C. Observation of entanglement between a single trapped atom and a single photon. *Nature* **428**, 153–157 (2004).
18. Bergamini, S. *et al.* Holographic generation of microtrap arrays for single atoms by use of a programmable phase modulator. *J. Opt. Soc. Am. B* **21**, 1889–1894 (2004).
19. Schlosser, N., Reymond, G., Protsenko, I. & Grangier, P. Sub-poissonian loading of single atoms in a microscopic dipole trap. *Nature* **411**, 1024–1027 (2001).
20. Dingjan, J. *et al.* A frequency-doubled, pulsed laser system for rubidium manipulation. *Appl. Phys. B* **82**, 47–51 (2006).
21. Darquié, B. *et al.* Controlled single-photon emission from a single trapped two-level atom. *Science* **309**, 454–456 (2005).
22. Legero, T., Wilk, T., Kuhn, A. & Rempe, G. Time-resolved two-photon quantum interference. *Appl. Phys. B* **77**, 797–802 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge support from the European Union through the Integrated Project 'SCALA'. J.D. was funded by Research Training Network 'CONQUEST'. M.P.A.J. was supported by a Marie Curie fellowship.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.B. (antoine.browaeys@iota.u-psud.fr).

Solution-processed silicon films and transistors

Tatsuya Shimoda^{1*}, Yasuo Matsuki^{2*}, Masahiro Furusawa^{1*}, Takashi Aoki¹, Ichio Yudasaka¹, Hideki Tanaka¹, Haruo Iwasawa², Daohai Wang², Masami Miyasaka¹ & Yasumasa Takeuchi^{2†}

The use of solution processes—as opposed to conventional vacuum processes and vapour-phase deposition—for the fabrication of electronic devices has received considerable attention for a wide range of applications^{1–7}, with a view to reducing processing costs. In particular, the ability to print semiconductor devices using liquid-phase materials could prove essential for some envisaged applications, such as large-area flexible displays. Recent research in this area has largely been focused on organic semiconductors^{8–11}, some of which have mobilities comparable to that of amorphous silicon¹¹ (a-Si); but issues of reliability remain. Solution processing of metal chalcogenide semiconductors to fabricate stable and high-performance transistors has also been reported^{12,13}. This class of materials is being explored as a possible substitute for silicon, given the complex and expensive manufacturing processes required to fabricate devices from the latter. However, if high-quality silicon films could be prepared by a solution process, this situation might change drastically. Here we demonstrate the solution processing of silicon thin-film transistors (TFTs) using a silane-based liquid precursor. Using this precursor, we have prepared polycrystalline silicon (poly-Si) films by both spin-coating and ink-jet printing, from which we fabricate TFTs with mobilities of $108\text{ cm}^2\text{ V}^{-1}\text{ s}^{-1}$ and $6.5\text{ cm}^2\text{ V}^{-1}\text{ s}^{-1}$, respectively. Although the processing conditions have yet to be optimized, these mobilities are already greater than those that have been achieved in solution-processed organic TFTs, and they exceed those of a-Si TFTs ($\leq 1\text{ cm}^2\text{ V}^{-1}\text{ s}^{-1}$).

We have pursued the development of a novel liquid precursor (herein referred to as ‘liquid silicon material’) that can be used in a liquid process to form a silicon film^{14,15}. Since the films produced from this precursor must be convertible to high purity silicon, potential candidates are limited to carbon- and oxygen-free hydrogenated silicon compounds. Typical hydrogenated silicon compounds are either of the straight-chain ($\text{Si}_n\text{H}_{2n+2}$) or cyclic (Si_nH_{2n}) forms. For $n \geq 3$, these compounds are liquid at room temperature and decompose to form a-Si when heated to 300°C or higher. However, for $n < 10$, boiling points are less than 300°C and such compounds evaporate before thermal decomposition, which makes solution processing difficult. Oligomeric and polymeric hydrogenated polysilanes, $-(\text{SiH}_2)_n-$, have received little attention since first being synthesized by the Kipping method^{16,17}, because of their limited solubility in organic solvents. Nevertheless, hydrogenated polysilanes are potentially ideal liquid silicon materials provided a suitable solvent can be found. Since cyclic silanes are known to undergo ring-opening polymerization^{18,19}, this reaction can be used for the development of liquid silicon materials. We have applied photo-induced ring-opening polymerization to obtain pure hydrogenated polysilanes from purified hydrogenated cyclic silanes. Among the hydrogenated cyclic silanes we chose cyclopentasilane

(CPS), Si_5H_{10} , since it is relatively stable and exhibits a high photoreactivity upon irradiation with ultraviolet (UV) light.

Using the method developed by Hengge *et al.*^{20,21}, CPS monomer (a clear and colourless liquid under ambient conditions) was synthesized. It has a boiling point of 194°C and is soluble in most organic solvents. After purification, the CPS was exposed to 405 nm UV light to induce photo-polymerization. During exposure, the CPS gradually became cloudy and viscous. After sufficient exposure to the UV source, the liquid was transformed into a white solid, presumably made up of a mixture of hydrogenated polysilanes. Although this solid is insoluble in all common organic solvents, it proved to be soluble in the CPS monomer precursor. The hydrogenated polysilane was also found to dissolve in a mixture of CPS and an organic solvent. This solvent mixture plays an important role in controlling the wettability, coating properties, and thickness of the resulting silicon films, all of which are difficult to control when CPS is used alone. In the actual process, UV irradiation is halted before the CPS completely polymerizes such that polysilanes of various molecular weights are dissolved in unreacted CPS. By diluting the solution with an organic solvent and then filtering out those insoluble polysilanes that precipitated as a result of dilution, we obtained the solution that we refer to as liquid silicon material.

The polymerization process of CPS was investigated using gel permeation chromatography (GPC). Figure 1 shows the results of the GPC measurements of the CPS (Fig. 1a) and the UV-irradiated CPS (Fig. 1b). Peaks corresponding to CPS and toluene were observed in both the non-irradiated and irradiated samples. In addition, in Fig. 1b, a broad peak was also observed around $M_w = 2,600$. This peak corresponds to polysilanes of various molecular weights, indicating CPS polymerization. The GPC measurements were also used to control and optimize the molecular weight distribution of the polysilanes which significantly affects the wettability of the precursor solution to a glass substrate.

The next step is to form an a-Si film by heating the spin-coated polysilane film in order to induce thermal decomposition. Three samples were used to study the thermal decomposition of the polysilane, through the use of thermal desorption spectroscopy (TDS; see Methods). The relationship between the preheating condition of polysilane and the amount of relevant gas (H_2 , SiH_2 and SiH_3) desorbed during the post-annealing phase of TDS are shown in Fig. 2. Sample a, prebaked at 300°C for 10 min, reveals an SiH_2 and SiH_3 desorption peak at around 280°C , followed by intensive hydrogen desorption between 300°C and 400°C , indicating that the polysilane is not completely converted to a-Si when baked at 300°C for 10 min. Given the binding energies of Si–Si (224 kJ mol^{-1}) and Si–H (318 kJ mol^{-1})²², the Si–Si bonds in the polysilane break first at a temperature lower than 280°C , followed next by the breaking of the Si–H bonds, where a three-dimensional silicon network starts to form. In sample b, prebaked at 300°C for 2 h,

¹Technology Platform Research Centre, Seiko Epson Corporation, 281 Fujimi, Fujimi-machi, Nagano-ken, 399-0293 Japan. ²Fine Electronic Research Laboratories, JSR Corporation, 100 Kawajiri-cho, Yokkaichi, Mie, 510-8552 Japan. [†]Present address: International Centre for Material Research, 1-1 Minamiwataridacho, Kawasaki-ku, Kawasaki-city, Kanagawa, 210-0855 Japan.

*These authors contributed equally to this work.

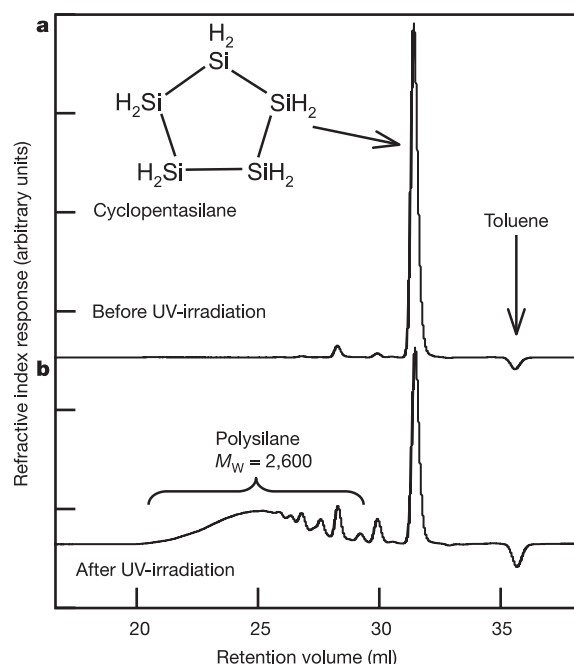


Figure 1 | Gel permeation chromatogram (GPC) of liquid precursor for Si film. **a**, Cyclopentasilane (CPS) and **b**, UV-irradiated CPS, both of which were diluted with toluene (1 vol.%) before GPC measurements. The UV-irradiation conditions were 405 nm, 100 mW cm⁻², and a 10 min irradiation for 1 cm³ of CPS. The broad peak around $M_w = 2,600$ corresponds to polysilanes of various molecular weights as a result of the photo-induced polymerization of CPS.

SiH₂ and SiH₃ desorption is three or four orders of magnitude lower than that of sample a. Sample c released much less desorption gas than samples a and b, since it had almost completely converted to a-Si during preheating at 540 °C for 2 h. Using TDS analysis, we estimate the total amount of hydrogen atoms (the atomic ratio of H/Si) in samples a, b and c, before TDS measurement, to be more than 22%, 3% and 0.3%, respectively. Thus the formation process of a-Si is inferred to be as follows: as the spin-coated polysilane film is heated, volatile components such as toluene and CPS evaporate first. Next, the Si-Si bonds in the polysilane begin to break at a temperature below 280 °C and a fraction of the polysilane is released as SiH₂ and SiH₃. After this, at around 300 °C, the Si-H bonds break, resulting in a three-dimensional a-Si network.

In order to test the electrical properties of the film, we fabricated simplified bottom gate TFTs using the spin-coated a-Si. The maximum mobility of these TFTs was 10⁻³–10⁻⁴ cm² V⁻¹ s⁻¹, which is three or four orders of magnitude smaller than that obtained through conventional plasma enhanced chemical vapour deposition (CVD) techniques. Such poor mobility is attributed to the low concentration of hydrogen atoms that terminate the dangling bonds in the film. Unlike a-Si film formed from conventional CVD, which contains 5–20% hydrogen, the spin-coated a-Si film baked at 540 °C for 2 h contains 0.3% hydrogen and appreciable quantities of dangling bonds that hinder the mobility. In contrast, film that has been spin-coated and baked at a temperature of 300 °C or less contains more than 20% hydrogen and could function somewhat as a semiconductor. However, film processed at such a low temperature might no longer be described as amorphous silicon and is easily oxidized in the air. Some kind of technical solution must be found in order to strike a balance between a lower process temperature and preventing oxidation.

To this end, we have studied the crystallization of the film (before optimization of the low-temperature a-Si process) in order to demonstrate the fundamental potential of this solution processing

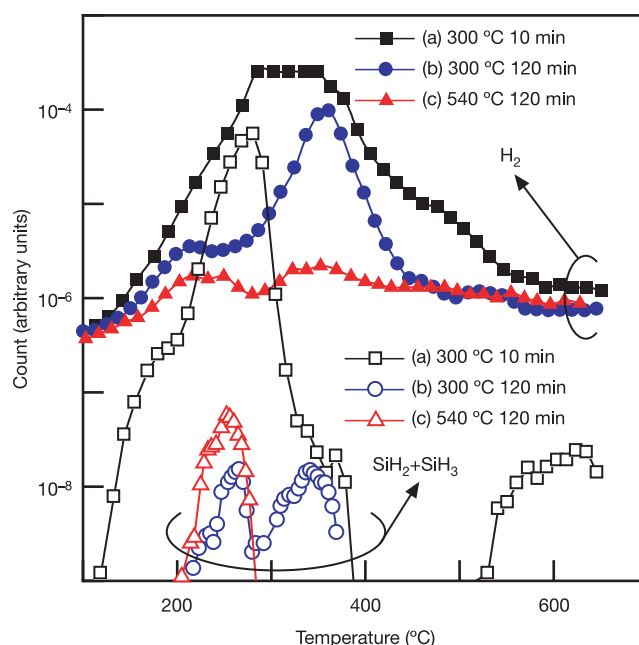


Figure 2 | Thermal desorption spectrum (TDS) of solution-processed a-Si film. Three samples were prepared by the thermal decomposition of polysilane under the following conditions: sample a, 300 °C for 10 min; sample b, 300 °C for 120 min; sample c, 540 °C for 120 min. Desorbed gases from the samples were analysed by mass spectroscopy while the samples were heated in a vacuum.

approach for low-temperature poly-Si (LTPS) TFT applications. In addition, low-hydrogen a-Si films are suitable for excimer laser crystallization, which is the standard crystallization method adopted during the commercial production of LTPS TFTs²³. The a-Si films, formed by coating the liquid silicon material and baking it at 540 °C, were irradiated by a 308 nm XeCl excimer laser at various laser energies. As the laser energy increases, the colour of the films changes from light auburn to light yellow, suggesting a conversion from the amorphous to the polycrystalline phase. The TEM image in Fig. 3 clearly shows the crystalline growth as a result of laser irradiation. Using Raman spectroscopy²⁴, this crystallization process was confirmed to be almost identical to that of a-Si fabricated through conventional CVD. The full-width at half-maximum of the crystalline peak in the Raman spectrum decreases sharply as the laser energy increases, reaching a minimum value of 6.3 cm⁻¹ at around 300 mJ cm⁻²; thereafter the peak slightly broadens, reflecting microcrystallization.

Next, we fabricated TFTs using the coating-formed poly-Si films followed by standard fabrication steps used for conventional LTPS TFTs (see Methods). As shown in Fig. 4a, these TFTs exhibit good electrical characteristics with field-effect mobility (calculated from the transconductance in the saturation region) ranging from 74 to 108 cm² V⁻¹ s⁻¹ in 15 transistors randomly selected among a 4-inch substrate. The transistor of mobility 108 cm² V⁻¹ s⁻¹, whose output characteristics are shown in Fig. 4b, also possesses an on/off ratio of seven digits, 5.0 V in threshold voltage V_{th} and 0.83 V per decade in s-factor (the gate voltage that induces a tenfold increase of drain current in the sub-threshold region).

The mobility was strongly affected not only by the conditions of laser crystallization but also by the amount of oxygen in the silicon film. In the poly-Si film that exhibited a mobility of 108 cm² V⁻¹ s⁻¹ the oxygen concentration was 1,100 p.p.m. Since the CPS and polysilanes are fairly oxygen-sensitive²⁵, several precautions had to be taken during the preparation and treatment of a liquid precursor. The coarsely synthesized CPS was distilled repeatedly to remove impurities including oxides before photo-polymerization.

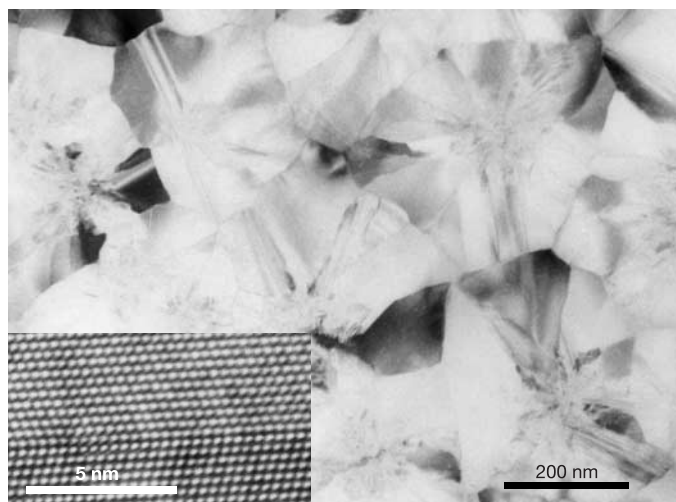


Figure 3 | A TEM image of a solution-processed poly-Si film. The film was formed by spin-coating and baking of the liquid silicon materials followed by laser crystallization. The high-resolution TEM image inserted in the figure clearly highlights the atomic image of the silicon crystal. The micrograph also illustrates that the grain size in the film is about 300 nm, which is comparable to that of conventional CVD-formed poly-Si film.

Furthermore, the organic diluent such as toluene was carefully deoxidized by leaving it in a dry box for several days before use. The oxygen concentration in the box was maintained strictly below 0.5 p.p.m. during all processes from spin-coating to baking. Such precautions enabled control of the oxygen content of the film to less than 2,000 p.p.m. The TFTs resulting from such low-oxygen films exhibited mobility as high as $50\text{--}100\text{ cm}^2\text{ V}^{-1}\text{ s}^{-1}$, while no significant correlations were observed between the oxygen concentration and mobility since manufacturing variations such as the laser condition are the dominant factors. However, a slight failure in controlling the oxygen level in the dry box yielded a silicon film with an oxygen concentration of 8,000 p.p.m. and TFTs with a maximum mobility of $20\text{ cm}^2\text{ V}^{-1}\text{ s}^{-1}$. By intentionally maintaining the oxygen level in the dry box at 10 p.p.m. during all processing, we obtained an insulating a-Si film that contained 7% oxygen.

Finally, we demonstrate the printing applicability of the liquid silicon material by fabricating TFTs using ink-jet formation of poly-Si islands (see Methods). The silicon island can be seen in the SEM image of Fig. 4c as the disk-shaped element with a rough surface. The nature of the external disk-shaped phase with a smooth surface is still unknown, but is presumably the residual remains of the droplet, formed during the drying and shrinking processes, which was once dispersed over the outer disk region. Since the wettability of the liquid silicon material was not sufficiently well understood to enable control of the shrinking behaviour of an ink droplet, the resulting a-Si island became too thick for laser crystallization. A thicker film generally requires a more intense laser for optimal crystallization. The actual laser intensities used for the crystallization of 60 nm spin-coated and 300 nm ink-jet-treated Si film were 345 mJ cm^{-2} and 450 mJ cm^{-2} , respectively. The latter intensity is not optimized but determined roughly from the sample thickness. With such an intense laser source, surface roughness is easily induced, as shown in Fig. 4c.

The ink-jet-printed (or 'ink-jetted') TFT operated with a mobility of $6.5\text{ cm}^2\text{ V}^{-1}\text{ s}^{-1}$ and an on/off ratio of three digits (Fig. 4a). This low mobility is attributed to the poor crystallinity and rough surface, while the large off current, which was confirmed to be the current between source and drain rather than a leak current through the gate insulator, is also attributed to the thick silicon film. Such a thick film contains many dangling bonds and defects near the substrate where crystallization is incomplete. The formation of thin, uniform a-Si film using an ink-jet process would improve the surface morphology

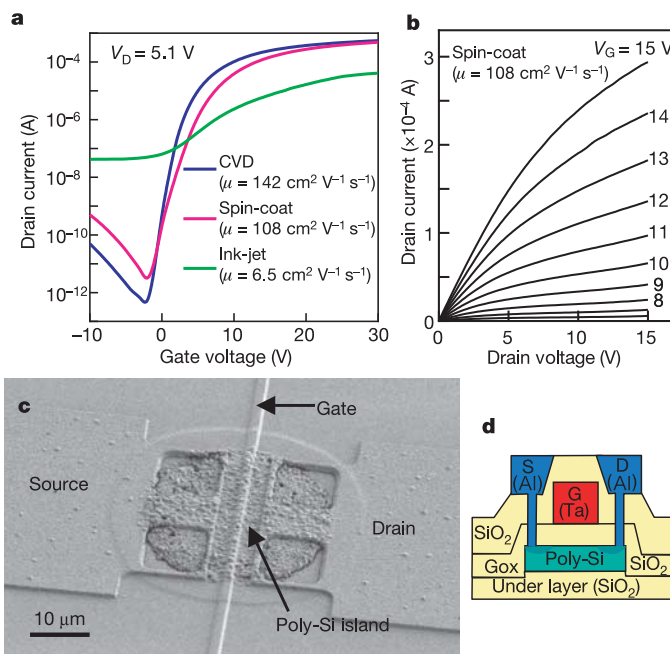


Figure 4 | The structure and characteristics of solution-processed LTPS TFTs. **a**, The transfer characteristics of LTPS TFTs, whose silicon film was formed by CVD (blue), spin-coating (magenta) and ink-jet printing (green), respectively. The drain current of the ink-jetted TFT is normalized to have the same channel width and length as the CVD-formed and spin-coated TFTs, for comparison. **b**, The output characteristics of the TFT using spin-coated silicon film whose transfer properties are shown in **a**. **c**, An SEM image of the TFT made from ink-jetted silicon film, whose transfer characteristics are shown in **a**. **d**, A cross-sectional schematic of a fabricated TFT. Gox, SiO_2 gate insulator.

and electrical properties of the poly-Si films. However, the wettability and behaviour of microdroplets are known to differ significantly from those of macroscopic droplets^{26,27}. A thorough understanding of microdroplets will be needed to realize the ink-jet printing of liquid silicon material for practical applications.

We have demonstrated that high-quality poly-Si film can be formed by spin-coating or ink-jetting liquid silicon material. The mobility of the spin-coated TFT, $108\text{ cm}^2\text{ V}^{-1}\text{ s}^{-1}$, is definitely that of an LTPS TFT and cannot be obtained with a-Si TFTs. Though currently inferior to that of conventional spin-coated LTPS TFTs, the performance of our ink-jetted TFT will probably improve with further advances in both the materials and processes employed. The ultimate goal of this research is to fabricate high-performance silicon TFTs by means of an all-liquid process, whereby all layers are directly patterned with liquid materials. To this end, we are currently working to develop liquid materials for films other than channel silicon, such as dielectric layers, doped silicon for source and drain regions, and metallic films for electrodes. The ink-jet process for these materials is also being studied. Such efforts are expected to establish a novel, low-energy, low-cost and high-throughput process for the fabrication of high-performance TFTs.

METHODS

Preparation and analysis of a-Si film. A 30 vol.% toluene solution of UV-irradiated CPS was spin-coated onto a quartz substrate to yield an approximately 100-nm-thick a-Si film after heat treatment. Heat treatment conditions (achieved using a hot plate) for samples a, b and c were 300°C for 10 min, 300°C for 120 min and 540°C for 120 min, respectively. All experiments were carried out in a nitrogen-filled dry box with a residual oxygen concentration of less than 0.5 p.p.m., and the oxygen level was monitored using a galvanic fuel cell sensor. Raman scattering spectra of these samples were recorded in order to confirm the a-Si nature of the films at the typical a-Si peak of around 480 cm^{-1} . Impurity concentrations in the films were investigated using secondary ion mass

spectroscopy. The results confirmed that the film was composed almost entirely of silicon, with only a few trace impurities. Alkali and alkali-earth metal impurities, which negatively affect TFT characteristics, were present at a level of less than 1 p.p.m. The carbon content in the resulting film was surprisingly low—only 200 p.p.m.—considering that an organic solvent was used as the starting material. By strictly controlling the oxygen content in the dry box to less than 0.5 p.p.m., we were able to limit the oxygen content of the film to less than 2,000 p.p.m. Thermal desorption spectroscopy (TDS) was used to investigate the process by which polysilane is converted to a-Si. In TDS, a sample is heated in a vacuum and the gases that are desorbed from the sample are analysed using mass spectroscopy.

Fabrication of TFTs by spin-coating. The n-channel TFTs, whose structure is schematically illustrated in Fig. 4d, were fabricated as follows. First, an SiO₂ underlayer was formed by plasma CVD on a quartz substrate. After cleaning the substrate surface by 172 nm UV-irradiation at 10 mW cm⁻² for 10 min, the liquid silicon material—12 vol.% toluene solution of UV-irradiated CPS—was spin-coated at 2,000 r.p.m. in a nitrogen-filled dry box. The spin-coated substrate was immediately placed on a hot plate heated to 200 °C and the temperature was raised to 400 °C within 10 min. After maintaining a temperature of 400 °C for 30 min, it was further increased to 540 °C over another 10 min period and held for 2 h to form a 50-nm-thick a-Si film. Next, the amorphous film was converted to a polycrystalline one by irradiating it with 308 nm wavelength excimer laser light at 345 mJ cm⁻². After the poly-Si had been etched to create islands, a 120-nm-thick SiO₂ gate insulator (Gox in Fig. 4d) was formed by plasma CVD, followed by tantalum sputtering and etching to form gate electrodes. The source and drain regions were formed by the self-aligned ion implantation of phosphorous ions using the gate electrodes as a mask. Finally, we made the TFTs accessible for measurement by forming an interlayer insulator, opening contact holes in the insulator to reveal the source and drain regions, and then sputtering aluminium to form electrodes. The channel width and length of the TFTs were both 10 µm. The coating-formed silicon film did not present any notable problems or difficulties during the above fabrication steps, as the silicon film was of semiconductor-grade purity and the film processing conditions—etching rate, laser conditions and so on—were nearly the same as those used for CVD-produced silicon film. For comparison, conventional TFTs were fabricated by the same process except that the silicon film was formed by CVD.

Fabrication of TFTs by ink-jet printing. TFTs with the same structure as above were formed using ink-jet printing instead of photolithography to form the channel silicon island. A 10 vol.% toluene solution of UV-irradiated CPS was suitable for ejection from a piezo-driven print head¹⁷. Since its viscosity is almost the same as that of toluene and its stability is such that it can be kept for months in a dark place at room temperature, the ejection of the solution from the print head was stable and reproducible. Three droplets, each of weight 10 ng, were deposited in the location where the channel island was to be formed. The droplets were converted into a poly-Si island of diameter 30–40 µm by baking at 540 °C (using the same steps as for spin-coated film) followed by 308 nm excimer laser crystallization at 450 mJ cm⁻². The island was 300 nm thick at the centre and became thinner towards the periphery of the sample. After the gate insulator had been formed, the same steps were applied as for the spin-coated TFTs. The channel width and length of the resulting TFTs were 36 µm and 2 µm, respectively.

Received 20 September 2005; accepted 23 January 2006.

1. Shimoda, T. *et al.* Multicolor pixel patterning of light-emitting polymers by ink-jet printing. In *1999 SID International Symposium Digest of Technical Papers* 376–379 (Society for Information Display, San Jose, 1999).
2. Miyashita, S. *et al.* Full color displays fabricated by ink-jet printing. In *Proc. 21st International Display Research Conference in Conjunction with 8th International Display Workshop (Asia Display / IDW '01)* 1399–1402 (2001).
3. Peumans, P., Uchida, S. & Forrest, S. R. Efficient bulk heterojunction photovoltaic cells using small-molecular-weight organic thin films. *Nature* **425**, 158–162 (2003).

4. Okamura, S., Takeuchi, R. & Shiozaki, T. Fabrication of ferroelectric Pb(Zr,Ti)₃ thin films with various Zr/Ti ratios by ink-jet printing. *Jpn. J. Appl. Phys.* **41**, 6714–6717 (2002).
5. Tahar, R. B. H., Ban, T., Ohya, Y. & Takahashi, Y. Optical, structural, and electrical properties of indium oxide thin films prepared by the sol-gel method. *J. Appl. Phys.* **82**, 865–870 (1997).
6. Yudasaka, I., Tanaka, H., Miyasaka, M., Inoue, S. & Shimoda, T. Poly-Si thin-film transistors using polysilazane-based spin-on glass for all dielectric layers. In *2004 SID International Symposium Digest of Technical Papers* 964–967 (Society for Information Display, San Jose, 2004).
7. Furusawa, M. *et al.* Inkjet-printed bus and address electrodes for plasma display. In *2002 SID International Symposium Digest of Technical Papers* 753–755 (Society for Information Display, San Jose, 2002).
8. Sirringhaus, H. *et al.* High-resolution inkjet printing of all-polymer transistor circuits. *Science* **290**, 2123–2126 (2000).
9. Kawase, T., Sirringhaus, H., Friend, R. H. & Shimoda, T. All-polymer thin film transistors fabricated by high-resolution inkjet printing. In *2000 International Electron Device Meeting (IEDM) Tech. Digest* 623–626 (2000).
10. Gelinck, G. H. *et al.* Flexible active-matrix displays and shift registers based on solution-processed organic transistors. *Nature Mater.* **3**, 106–110 (2004).
11. Afzali, A., Dimitrakopoulos, C. D. & Breen, T. L. High-performance, solution-processed organic thin film transistors from a novel pentacene precursor. *J. Am. Chem. Soc.* **124**, 8812–8813 (2002).
12. Ridley, B. A., Nivi, B. & Jacobson, J. M. All-inorganic field effect transistors fabricated by printing. *Science* **286**, 746–749 (1999).
13. Mitzi, D. B., Kosbar, L. L., Murray, C. E., Copel, M. & Afzali, A. High-mobility ultrathin semiconducting films prepared by spin coating. *Nature* **428**, 299–303 (2004).
14. Aoki, T. *et al.* Method of manufacturing device, device, and electronic apparatus. US Patent Application 0029364 (2004).
15. Shimoda, T. *et al.* Method for forming silicon film. US Patent 6541354 (2003).
16. Kipping, F. S. Organic derivatives of silicon. Complex silicohydrocarbons [SiPh₂]_n. *J. Chem. Soc.* **125**, 2291–2297 (1924).
17. John, P., Oder, I. M. & Wood, J. The electrical conductivity of polysilane, (SiH₂)_x. *J. Chem. Soc. Chem. Commun.* 1496–1497 (1983).
18. Suzuki, M., Kotani, J., Gyobu, S., Kaneko, T. & Saegusa, T. Synthesis of sequence-ordered polysilane by anionic ring-opening polymerization of phenylnonamethycyclopentasilane. *Macromolecules* **27**, 2360–2363 (1994).
19. Cypryk, M., Gupta, Y. & Matyjaszewski, K. Anionic ring-opening polymerization of 1,2,3,4-tetramethyl-1,2,3,4-tetraphenylcyclotetrasilane. *J. Am. Chem. Soc.* **113**, 1046–1047 (1991).
20. Hengge, E. & Bauer, G. Cyclopentasilan, das erste unsubstituierte cyclische Siliciumhydrid. *Angew. Chem.* **85**, 304–305 (1973).
21. Hengge, E. & Bauer, G. Darstellung und Eigenschaften von Cyclopentasilan. *Monatsh. Chem.* **106**, 503–512 (1975).
22. Raabe, G. & Michl, J. Multiple bonding to silicon. *Chem. Rev.* **85**, 419–509 (1985).
23. Sameshima, T., Usui, S. & Sekiya, M. XeCl excimer laser annealing used in the fabrication of poly-Si TFTs. *IEEE Electron Device Lett.* **7**, 276–278 (1986).
24. Kitahara, K., Yamazaki, R., Kurosawa, T., Nakajima, K. & Moritani, A. Analysis of stress in laser-crystallized polysilicon thin films by Raman scattering spectroscopy. *Jpn. J. Appl. Phys.* **41**, 5055–5059 (2002).
25. Chatgililoglu, C. *et al.* Autoxidation of poly(hydrosilane)s. *Organometallics* **17**, 2169–2176 (1998).
26. Morii, K. *et al.* Characterization of light-emitting polymer devices prepared by ink-jet printing. In *Proc. 10th Int. Workshop on Inorganic and Organic Electroluminescence* 357–360 (2000).
27. Morii, K., Masuda, T., Ishida, M., Hotta, S. & Shimoda, T. The direct patterning of crystalline organic-semiconductor films on a substrate by ink-jet printing. In *Proc. Int. Conference on Synthetic Metals* 126–127 (2004).

Acknowledgements We thank the members of Seiko Epson Corporation's pilot line and ink-jet industrial application project, for fabricating TFTs and ink-jet experiments in this research. This work is partially supported by a grant from the New Energy and Industrial Technology Development Organization (NEDO).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.F. (furusawa.masahiro@exc.epson.co.jp).

Increased Arctic cloud longwave emissivity associated with pollution from mid-latitudes

Timothy J. Garrett¹ & Chuanfeng Zhao¹

There is consensus among climate models that Arctic climate is particularly sensitive to anthropogenic greenhouse gases and that, over the next century, Arctic surface temperatures are projected to rise at a rate about twice the global mean¹. The response of Arctic surface temperatures to greenhouse gas thermal emission is modified by Northern Hemisphere synoptic meteorology and local radiative processes^{2–4}. Aerosols may play a contributing factor through changes to cloud radiative properties. Here we evaluate a previously suggested contribution of anthropogenic aerosols to cloud emission and surface temperatures in the Arctic^{5–8}. Using four years of ground-based aerosol and radiation measurements obtained near Barrow, Alaska, we show that, where thin water clouds and pollution are coincident, there is an increase in cloud longwave emissivity resulting from elevated haze levels. This results in an estimated surface warming under cloudy skies of between 3.3 and 5.2 W m^{−2} or 1 and 1.6 °C. Arctic climate is closely tied to cloud longwave emission^{2,4,9}, but feedback mechanisms in the system are complex¹⁰ and the actual climate response to the described sensitivity remains to be evaluated.

Previous studies of the effects of anthropogenic aerosols on cloud radiative properties have emphasized their potential contributions to planetary cooling, through greater activation of cloud droplets, and an accompanying increase in the reflection to space of solar energy^{1,11,12}. Large, soluble aerosols act as cloud condensation nuclei (CCN). In what is known as the ‘first indirect effect’¹¹, for a fixed water path W , pollution increases shortwave (SW) cloud albedo α because higher concentrations of CCN correspond to smaller droplets with higher number concentrations N , and hence higher cloud scattering cross-sections. This sensitivity can be quantified through $S_{SW} = (d\alpha/dN)_W$. In pristine stratus clouds, S_{SW} might range as high as 0.01 cm³, sufficient to partly offset greenhouse gas warming^{1,12}.

Our goal here is to evaluate a hypothesized contribution to cloud longwave (LW) emission and surface temperatures from the long-range transport to the Arctic of mid-latitude anthropogenic aerosols^{5–7}. Cloud thermal emission follows the relation $F_{LW} = \varepsilon \sigma_{SB} T^4$, where ε is the emissivity, σ_{SB} the Stefan–Boltzmann constant and T temperature. Emissivity is related to cloud properties through $\varepsilon = 1 - \exp(-kW)$, where $W = 4/3\pi\rho r^3Nh$ (ρ being the bulk density of water, r the mean droplet radius, and h the cloud thickness), and k depends on the wavelength and cloud droplet size. Traditionally, the relationship between CCN concentrations and ε has not been a focus, because it has often been assumed that values of W in water clouds are greater than about 50 g m^{−2}, in which case they approximate a blackbody ($\varepsilon \approx 1$); accordingly, ε is insensitive to changes in cloud microstructures. However, in recent years, better detection techniques have increased awareness of thin water clouds¹³. As greybodies, with $\varepsilon < 1$, it is hypothesized that the emissivity of these clouds may be linked to changes in droplet concentration⁷. From estimates of k (ref. 7), it is straightforward to show that the LW sensitivity of ε to N , $S_{LW} = (d\varepsilon/dN)_W$ is given by Fig. 1.

S_{LW} exceeds 0.01 cm³ in pristine clouds with $N < 10$ cm^{−3} and $W < 25$ g m^{−2}.

In the Arctic, this LW indirect effect is relevant because the surface radiation balance between winter and early spring is tied to downwelling thermal fluxes from thin, greybody, low to mid-level clouds^{14,15}. At this time of year, the Arctic is characterized by widespread pollution known as Arctic haze. Strong east–west pressure gradients cause episodic northward intrusions of polluted air from mid-latitude Eurasia and North America^{16,17}. Because precipitation is low, aerosol lifetimes are long¹⁸, and pollution accumulates through the depth of the lower troposphere¹⁹. Arctic haze only disperses when moist destabilizing air intrudes in spring²⁰. Indirect solar forcing by the haze aerosol is usually small or negligible because the polluted months are either dark, or characterized by high solar zenith angles and a surface that is already bright^{14,15}. Instead, clouds are thin, so there is significant potential for increased surface warming from aerosol modifications to cloud LW emissivity⁷.

To examine this more closely, we use long-term ground-based data sets from two adjacent sites near Barrow, Alaska, supervised by the DOE Atmospheric Radiation Measurement (ARM) programme²¹,

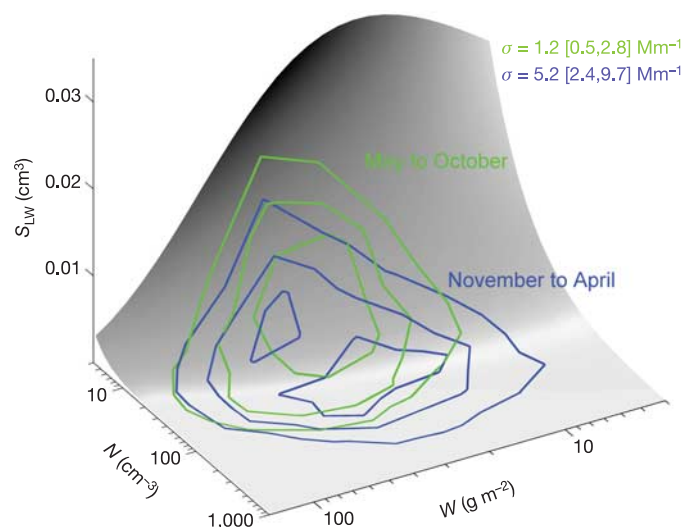


Figure 1 | Sensitivity of cloud LW emissivity to changes in droplet number concentration for fixed water path, $S_{LW} = (d\varepsilon/dN)_W$. The surface is theoretically based, assuming a cloud depth and temperature of 100 m and 253 K, respectively. Superimposed contours show quartiles in distributions of ground-based retrievals²⁷ of low-level water cloud number concentration N and liquid water path W from near Barrow, Alaska, between 2000 and 2003. Relative pollution levels for the contours are indicated by mean values of σ with [lower, upper] quartiles (σ is the 550 nm light scattering cross-section density of effloresced aerosol particles).

¹Department of Meteorology, University of Utah, Salt Lake City, Utah 84112, USA.

and by the NOAA Climate Monitoring Diagnostics Laboratory (CMDL)²². No measurements were made of CCN at either site. Instead, relative pollution levels are inferred from CMDL measurements of σ , the 550 nm light scattering cross-section density of effloresced aerosol particles. Usually CCN and σ correlate closely^{23,24} because the size range of aerosol normally associated with CCN is also highly efficient at scattering light. Measurements are limited to particles smaller than 1 μm . This is done to exclude natural contributions from larger Asian dust aerosol, which can initiate a conversion to ice in super-cooled water clouds²⁵. Analyses are further limited to time periods when the CMDL was upwind of Barrow. Retrievals of cloud ε , W , r_e (the area-weighted mean—or effective—radius of the droplet size distribution) and N are obtained using a ground-based Fourier transform infrared spectrometer (FTIR), lidar, radar, and temperature and humidity soundings from ARM, together with Global Ozone Monitoring Experiment (GOME) stratospheric ozone profiles from aboard the European Space Agency's Second European Remote Sensing Satellite (ERS-2)²⁶. The retrieval method is based on one developed previously for application to ice clouds in the Antarctic²⁷. To reduce uncertainty associated with vertical stratification when comparing surface aerosol measurements to retrieved cloud properties¹⁹, we restrict analyses to single-layer clouds with tops below 1.5 km altitude. Following application of the above restrictions, 9,440 5-min cloud samples were examined for the years between 2000 and 2003.

During this period, median retrieved droplet concentrations in greybody Arctic stratus were higher between November and April (56 cm^{-3} , 3,040 samples) than in the less polluted May to October (27 cm^{-3} , 6,400 samples) (Fig. 1). The late spring to autumn values are consistent with *in situ* measurements from the pristine central Arctic in summer²⁸. We then estimate from Fig. 1 that the LW sensitivity of clean Arctic stratus to changes in CCN concentration is $S_{\text{LW}} = 0.002 \text{ cm}^3$. We may assess the potential for cloud emissivity perturbation due to long-range transport of pollution as follows: clouds important to the Arctic LW surface radiation balance have bases typically below about 4 km altitude¹⁵. To raise concentrations of typical haze aerosol that might act as CCN¹⁶ to this depth by 1 cm^{-3} over the entire Arctic requires about 1 kt of material. The characteristic lifetime of Arctic haze particles ranges up to 39 days (ref. 18). Accordingly, sustaining a uniform anthropogenic increase in CCN concentrations of about 5 cm^{-3} (the amount necessary for a nominal increase to ε of ~ 0.01) requires long range transport to the Arctic of between 10 and 100 kt of material per year. For comparison, this is just 1% of the $1\text{--}10 \text{ Mt yr}^{-1}$ anthropogenic SO_2 made available to the Arctic from the Eurasian region north of 60° latitude alone¹⁶. Therefore, mid-latitude pollution should be sufficient to alter Arctic cloud LW properties.

The sensitivity of low-level cloud emissivity to pollution near Barrow is shown in Fig. 2a. Retrievals of cloud ε are sorted according to upper (polluted) and lower (clean) quartile thresholds in σ . To isolate the sensitivity to N when ε is also sensitive to W , values of ε are grouped into four logarithmically spaced bins in W , according to the exponential relationship between ε and W . As an artefact of the binning, there are small differences in mean W that bias differences in the mean values of ε for each cohort, but by at most $+0.01$. As expected, ε is insensitive to pollution in clouds with $W > 40 \text{ g m}^{-2}$, because the clouds already approximate blackbodies. In thinner clouds, however, a one-sided Student t -test shows, to within 95% confidence ($t > 4.0$ in each case), that polluted clouds have higher average values of ε than clean clouds with similar mean W : $\Delta\varepsilon$ ranges from 0.05 for the $5\text{--}10 \text{ g m}^{-2}$ bin, to 0.08 for the $20\text{--}40 \text{ g m}^{-2}$ bin. The reason that higher values of greybody emissivity occurred in the nominally polluted clouds was that their droplets had, on average, values of r_e that were $3 \mu\text{m}$ smaller, and concentrations three times more numerous, than droplets in nominally clean clouds.

Measurements from the SHEBA (Surface Heat Budget of the Arctic) experiment show that LW surface radiative forcing CF_{LW} by blackbody

clouds is $65 \pm 10 \text{ W m}^{-2}$ (ref. 15). Long-term measurements from drifting ice stations show that, between November and March, cloudy sky CF_{LW} is between 20 and 30 W m^{-2} and that, following a winter-time transition from clear to cloudy skies, Arctic surface temperatures increase by 6 to 9 K over 1 to 2 days (ref. 14). From these observations, and the results shown in Fig. 2a, we estimate that the upper quartile in σ is associated with an average additional $3.3\text{--}5.2 \text{ W m}^{-2}$, or 1 to 1.6 K surface warming by clouds, compared to the lower quartile in σ . Such sensitivity is consistent with previous estimates²⁴, based on more theoretical considerations, which predicted that a doubling in σ would increase Arctic CF_{LW} by 2 W m^{-2} . More recently, pyrgeometer data were used to attribute an average 8.2 W m^{-2} increased surface LW flux to aerosols from industrial emissions, 3.4 W m^{-2} of which was from changes to cloud r_e alone⁸.

An added consideration to our analysis is that, to be important to the Arctic surface radiation balance, clouds and pollution must coincide. Near Barrow, pollution levels are in the upper quartile of the entire data set at least one-half of the time between December and March (Fig. 2b). During this period, single-layer liquid clouds with bases below 4 km are normally greybodies. Their coverage, A , is less than in summer, but is nonetheless nearly one-half. It is difficult to compare simultaneous measurements of σ at the ground with cloud properties well aloft. However, our data set showed that high pollution levels were present 60% of the time when low-level water clouds with tops $< 1.5 \text{ km}$ were present. More representative of the Arctic as a whole, ice station data show that A varies between 0.5 and 0.6 in winter¹⁴, slightly more often than near Barrow.

We note that smaller additional LW aerosol forcing might be expected from pollution events in the two middle quartiles of σ . Further, it is possible that LW aerosol forcing may be modified by feedbacks that increase or decrease cloudy thermal emission through changes to cloud W and A . During the darker months, Arctic stratus

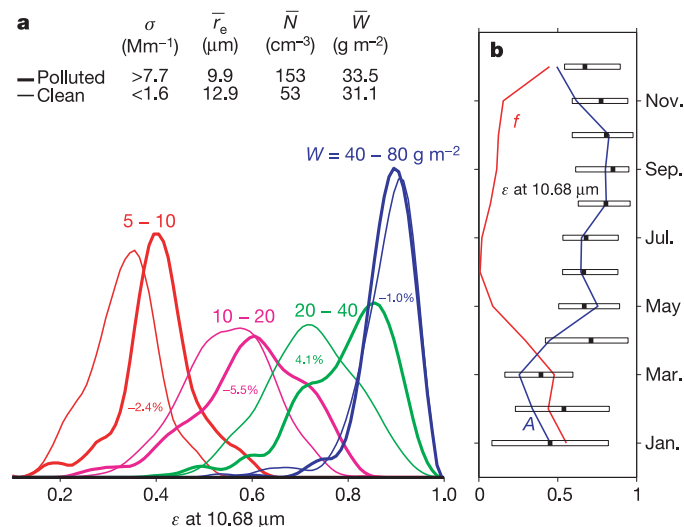


Figure 2 | Surface-based observations of Arctic pollution, clouds, and their radiative properties, made near Barrow, Alaska. **a**, Restricting cloud tops to below 1.5 km, we show probability distribution functions of values of liquid cloud emissivity ε within an atmospheric window at $10.68 \mu\text{m}$ wavelength. Values are sorted according to the upper (polluted) and lower (clean) quartile thresholds in surface measurements of σ , and binned according to retrieved values of water path, W . Differences in mean W between each pollution cohort are shown in per cent. Mean cloud properties shown are for greybody clouds only (r_e , area-weighted effective radius of the droplet size distribution; N , number concentration of droplets). **b**, For single-layer liquid clouds with bases less than 4 km (those radiatively important to the surface LW balance¹⁵ and for which cloud properties could be retrieved), we show the monthly-averaged cycle in ε (shown as a quartile plot), cloud coverage A , and the fraction of time f that σ is in the upper quartile for the entire data set.

is largely decoupled from the surface, and it maintains itself through cloud-top thermal emission²⁹. Although aerosol–cloud emissivity feedbacks are poorly understood, they are likely to be positive in this case because cloud-top thermal emission is proportional to ϵ . Airborne observations of a transition from clean to polluted air in a summer Arctic stratus deck showed increasing N and decreasing r_e , but also a halt to precipitation, higher water contents, and higher cloud A (ref. 7).

METHODS

Greybody cloud r_e and W are inferred using a minimization technique so that they are theoretically consistent with (1) FTIR retrievals of cloud ϵ , centred at two frequencies (862.5 and 935.8 cm^{-1}) in the atmospheric window where gaseous emission is particularly small, and (2) cloud transmission of stratospheric ozone emission centred at 1,040 cm^{-1} (ref. 27). This method is particularly well suited to high latitudes, because clouds tend to be greybodies, continuum emission by water vapour is low, and equilibrium stratospheric ozone concentrations are high. We have made several minor modifications to this technique. Instead of CO_2 slicing, we use lidar, radar, and ARM atmospheric soundings to determine cloud base, top, and temperature and humidity profiles, respectively. Second, water rather than ice clouds are selected³⁰ because, at low levels, super-cooled water clouds predominate in the Arctic, even at temperatures well below freezing¹⁰. Errors in retrievals for greybody clouds with $W < 40 \text{ g m}^{-2}$ were estimated from comparisons with the retrieved properties of synthetic plane-parallel clouds, and from the model sensitivities to uncertainties in the measurements. Combined, estimated errors are to within 10% for r_e , W and ϵ , and 30% for N . Cloud droplet number concentration is related to r_e and W by $N = 3W \exp(3s^2)/4\pi r_e^3 h$, where s is an assumed width of a log-normal droplet size distribution (a value of 0.32 ± 0.1 is chosen here on the basis of aircraft measurements)²⁴.

Received 29 July 2005; accepted 3 February 2006.

- Houghton, J. T. *et al.* (eds) *Climate Change 2001: The Scientific Basis* (Cambridge Univ. Press, Cambridge, UK, 2001).
- Overland, J. E., Adams, J. M. & Bond, N. A. Regional variation of winter temperatures in the Arctic. *J. Clim.* **10**, 821–837 (1997).
- Wang, X. & Key, J. R. Recent trends in Arctic surface, cloud, and radiation properties from space. *Science* **299**, 1725–1728 (2003).
- Vavrus, S. The impact of cloud feedbacks on Arctic climate under greenhouse forcing. *J. Clim.* **17**, 603–615 (2004).
- Zhang, T., Stamnes, K. & Bowling, S. A. Impact of clouds on surface radiative fluxes and snowmelt in the Arctic and subarctic. *J. Clim.* **9**, 2110–2123 (1996).
- Curry, J. & Ebert, E. Sensitivity of the thickness of Arctic sea ice to the optical properties of clouds. *Ann. Glaciol.* **14**, 43–46 (1990).
- Garrett, T. J., Radke, L. F. & Hobbs, P. V. Aerosol effects on the cloud emissivity and surface longwave heating in the arctic. *J. Atmos. Sci.* **59**, 769–778 (2002).
- Lubin, D. & Vogelmann, A. M. A climatologically significant aerosol longwave indirect effect in the Arctic. *Nature* **439**, 453–456 (2006).
- Miller, J. R. & Russell, G. L. Projected impact of climate change on the energy budget of the Arctic Ocean by a global climate model. *J. Clim.* **15**, 3028–3042 (2002).
- Curry, J. A., Rossow, W. B., Randall, D. & Schramm, J. L. Overview of Arctic cloud and radiation characteristics. *J. Clim.* **9**, 1731–1764 (1996).
- Twomey, S. The influence of pollution on the shortwave albedo of clouds. *J. Atmos. Sci.* **34**, 1149–1152 (1977).
- Twomey, S. Aerosol, clouds, and radiation. *Atmos. Environ. A* **25**, 2435–2442 (1991).
- Marchand, R. *et al.* An assessment of microwave absorption models and retrievals of cloud liquid water using clear-sky data. *J. Geophys. Res.* **108**, 4773, doi:10.1029/2003JD003843 (2003).
- Walsh, J. E. & Chapman, W. L. Arctic cloud-radiation-temperature associations in observational data and atmospheric reanalyses. *J. Clim.* **11**, 3030–3045 (1998).
- Shupe, M. D. & Intrieri, J. M. Cloud radiative forcing of the Arctic surface: the influence of cloud properties, surface albedo, and solar zenith angle. *J. Clim.* **17**, 616–628 (2004).
- Barrie, L. A. Arctic air pollution: an overview of current knowledge. *Atmos. Environ.* **20**, 643–663 (1986).
- Polissar, A. V. *et al.* The aerosol at Barrow, Alaska: long-term trends and source locations. *Atmos. Environ.* **33**, 2441–2458 (1999).
- Baskaran, M. & Shaw, G. E. Residence time of arctic haze aerosols using the concentrations and activity ratios of ^{210}Po , ^{210}Pb and ^7Be . *J. Aerosol. Sci.* **32**, 443–452 (2001).
- Wylie, D. P. & Hudson, J. G. Effects of long-range transport and clouds on cloud condensation nuclei in the springtime Arctic. *J. Geophys. Res.* **107**, doi:10.1029/2001JD000759 (2002).
- Sirois, A. & Barrie, L. A. Arctic lower tropospheric aerosol trends and composition at Alert, Canada: 1980–1995. *J. Geophys. Res.* **104**, 11599–11618 (1999).
- Stamnes, K., Ellingson, R. G., Curry, J. A., Walsh, J. E. & Zak, B. D. Review of science issues, deployment strategy, and status for the ARM North Slope of Alaska-Adjacent Arctic Ocean climate research site. *J. Clim.* **12**, 46–63 (1999).
- NOAA CMDL Point Barrow Observatory. (<http://www.cmdl.noaa.gov/obop/BRW/>) (2005).
- Radke, L. F., Hobbs, P. V. & Pinnons, J. E. Observations of cloud condensation nuclei, sodium-containing particles, ice nuclei and the light-scattering coefficient near Barrow, Alaska. *J. Appl. Meteorol.* **15**, 982–995 (1976).
- Garrett, T. J., Zhao, C., Dong, X., Mace, G. G. & Hobbs, P. V. Effects of varying aerosol regimes on low-level Arctic stratus. *Geophys. Res. Lett.* **31**, doi:10.1029/2004GL019928 (2004).
- Sassen, K. Dusty ice clouds over Alaska. *Nature* **434**, 456 (2005).
- Burrows, J. P. *et al.* The Global Ozone Monitoring Experiment GOME: Mission concept and first scientific results. *J. Atmos. Sci.* **56**, 151–175 (1999).
- Mahesh, A., Walden, V. P. & Warren, S. G. Ground-based remote sensing of cloud properties over the Antarctic Plateau. Part II: Cloud optical depths and sizes. *J. Appl. Meteorol.* **40**, 1279–1294 (2001).
- Bigg, E. K. & Leck, C. Cloud-active particles over the central Arctic Ocean. *J. Geophys. Res.* **106**, 32155–32166 (2001).
- Morrison, H. & Pinto, J. O. Mesoscale modeling of springtime Arctic mixed-phase stratiform clouds using a new two-moment bulk microphysics scheme. *J. Atmos. Sci.* **62**, 3683–3704 (2005).
- King, M. D. *et al.* Remote sensing of liquid water and ice cloud optical thickness and effective radius in the Arctic: Application of airborne multispectral MAS data. *J. Atmos. Ocean. Technol.* **21**, 857–875 (2004).

Acknowledgements This work was supported by the National Science Foundation. J. Ogren provided CMDL data. V. Walden and A. Mahesh provided assistance with retrieval development.

Author Contributions Both authors contributed equally to this manuscript.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.J.G. (tgarett@met.utah.edu).

LETTERS

Deciphering the evolution and metabolism of an anammox bacterium from a community genome

Marc Strous¹, Eric Pelletier², Sophie Mangenot², Thomas Rattei³, Angelika Lehner^{4,5}, Michael W. Taylor⁴, Matthias Horn⁴, Holger Daims⁴, Delphine Bartol-Mavel², Patrick Wincker², Valérie Barbe², Nuria Fonknechten², David Vallenet², Béatrice Segurens², Chantal Schenowitz-Truong², Claudine Médigue², Astrid Collingro⁴, Berend Snel⁶, Bas E. Dutilh⁶, Huub J. M. Op den Camp¹, Chris van der Drift¹, Irina Cirpus⁷, Katinka T. van de Pas-Schoonen¹, Harry R. Harhangi¹, Laura van Niftrik^{1,7}, Markus Schmid¹, Jan Keltjens¹, Jack van de Vossenberg¹, Boran Kartal¹, Harald Meier⁸, Dmitrij Frishman³, Martijn A. Huynen⁶, Hans-Werner Mewes^{3,9}, Jean Weissenbach², Mike S. M. Jetten^{1,7}, Michael Wagner⁴ & Denis Le Paslier²

Anaerobic ammonium oxidation (anammox) has become a main focus in oceanography and wastewater treatment^{1,2}. It is also the nitrogen cycle's major remaining biochemical enigma. Among its features, the occurrence of hydrazine as a free intermediate of catabolism^{3,4}, the biosynthesis of ladderane lipids^{5,6} and the role of cytoplasm differentiation⁷ are unique in biology. Here we use environmental genomics^{8,9}—the reconstruction of genomic data directly from the environment—to assemble the genome of the uncultured anammox bacterium *Kuenenia stuttgartiensis*¹⁰ from a complex bioreactor community. The genome data illuminate the evolutionary history of the Planctomycetes and allow us to expose the genetic blueprint of the organism's special properties. Most significantly, we identified candidate genes responsible for ladderane biosynthesis and biological hydrazine metabolism, and discovered unexpected metabolic versatility.

Since the biological nitrogen cycle was drafted at the end of the nineteenth century, the possibility of anaerobic ammonium oxidation (anammox) was generally overlooked². A chance discovery just ten years ago led to the identification of the responsible chemolithoautotrophic bacteria and the appreciation of their applied and ecological significance^{2,10}. Currently it is estimated that anammox contributes up to 50% to the removal of fixed nitrogen from the oceans globally¹. Unfortunately, anammox bacteria divide only once in two weeks at maximum speed and they are not available in pure culture². Environmental genomics^{8,9} thus offers a unique possibility to gain insight into the metabolism and evolution of these important bacteria. A laboratory bioreactor, anoxic and fed with synthetic wastewater containing the substrates ammonium, nitrite and bicarbonate, was the starting point for the present study. The bioreactor had been operated in the laboratory for one year, corresponding to 10 to 15 generations of anammox bacteria. Fluorescence *in situ* hybridization with a suite of ribosomal-RNA-targeted probes showed that this reactor was dominated (73 ± 5%) by the anammox bacterium *K. stuttgartiensis*. No other known anammox bacterium was detectable¹⁰. Comparative analysis of all 16S rRNA gene sequences retrieved from the different libraries constructed for genome sequencing (see below) showed that the microbial community as a whole contained at least six recognized bacterial phyla and

two lineages made up exclusively of uncultured bacteria (Supplementary Fig. 1). In total, 29 different operational taxonomic units were detected.

Genomic DNA was extracted from the bioreactor, and shotgun, fosmid and bacterial artificial chromosome (BAC) libraries were generated. Scaffolds were constructed from the shotgun library after binning, and gaps were closed with BAC and fosmid clones as well as polymerase chain reaction. Ultimately, five contigs were assembled (Supplementary Materials and Methods, Supplementary Fig. 2 and Supplementary Table 1). The five remaining gaps could not be closed and their size remains unknown. The robustness of the assembly was thoroughly validated with more than 100 confirmation points (Supplementary Fig. 2). The near completeness and correct assembly of the *K. stuttgartiensis* genome was further confirmed by the lack of suspicious redundancy or missing essential genes in major biosynthetic pathways. Apart from a gene for leucyl-transfer-RNA synthetase, the 64 clusters of orthologous groups of proteins (COGs) present in all currently sequenced bacterial genomes represented in the STRING database¹¹ were also present here. We therefore estimate the genome to be more than 98% (1 – 1/64) complete. Single nucleotide polymorphisms were virtually absent from the assembled scaffolds (less than 1 per 15,000 bases), strongly suggesting that this enrichment culture contained a single anammox strain.

The almost complete, 4.2-megabase genome of *K. stuttgartiensis* could then be used to deduce the biochemical pathway of anaerobic ammonium oxidation. Anammox is the one-to-one combination of ammonium and nitrite into dinitrogen gas. It was already known that hydrazine (N₂H₄) is an intermediate, and that the bacteria oxidize hydrazine to dinitrogen gas with a hydroxylamine-oxidoreductase-like protein (HAO)^{2–4}. On the basis of this experimental evidence, a pathway for anammox catabolism was previously postulated and involves hydrazine and hydroxylamine (NH₂OH)³. In the *K. stuttgartiensis* genome we detected more than 200 genes relevant to anammox catabolism and respiration (Supplementary Table 2).

Two hundred genes directly involved in catabolism and respiration exceeds the number possessed by most other bacteria (Supplementary Fig. 3). So far, a comparable level of redundancy has only been observed for versatile heterotrophic bacteria such as *Geobacter*

¹Department of Microbiology, IWW, Radboud University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, the Netherlands. ²CNRS UMR 8030 and Genoscope, 2 rue Gaston Crémieux CP 5706, 91057 Evry, France. ³Department of Genome Oriented Bioinformatics, Technische Universität München, Am Forum 1, 85354 Freising, Germany.

⁴Department of Microbial Ecology, University of Vienna, Althanstr. 14, 1090 Vienna, Austria. ⁵Institut für Lebensmittelsicherheit und Hygiene, Universität Zürich, CH-8057 Zürich, Switzerland. ⁶Center for Molecular and Biomolecular Informatics, Nijmegen Center for Molecular Life Sciences, Radboud University Nijmegen Medical Centre, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands. ⁷Kluyver Laboratory for Biotechnology, Delft University of Technology, 2628 BC Delft, The Netherlands. ⁸Lehrstuhl für Rechnerarchitektur und Rechnerorganisation, Parallelrechnerarchitektur, Institut für Informatik, Technische Universität München, 85748 Garching, Germany. ⁹Institute for Bioinformatics, GSF-National Research Center for Environment and Health, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany.

sulfurreducens and *Shewanella oneidensis*^{12,13}, while the aerobic ammonia oxidizer *Nitrosomonas europaea* has only 50 such genes¹⁴. Multiple divergent paralogues of respiratory complexes indicate the presence of a branched respiratory chain¹⁵ which would enable the respiration of different energy sources with different electron acceptors. Triggered by the large number of encoded *c*-type cytochromes and the previous detection of many *c*-type cytochromes in the metal-respiring organisms *G. sulfurreducens* and *S. oneidensis*, we experimentally investigated iron and manganese respiration by *K. stuttgartiensis*. Indeed, both iron and manganese oxides were respired with formate as electron donor (Supplementary Table 3). Iron was also oxidized with nitrate as electron acceptor. Thus the present study confirms and extends the recently discovered versatility¹⁶ of anammox bacteria.

Among the proteins encoded by the above-mentioned 200 genes were nine HAO-like proteins and two enzymes active in denitrification, nitrite:nitrate oxidoreductase (NarGH) and *cd*₁ nitrite:nitric oxide oxidoreductase (NirS). The presence of NirS was not compatible with the previously postulated model for anammox catabolism³, but the new genomic data allowed the deduction of a new metabolic pathway for anammox involving NO as intermediate (Fig. 1a). This new pathway is the only possible one consistent with the available experimental data^{2–4}, is thermodynamically and biochemically feasible, and conforms to ‘Ockham’s razor’—that is, it invokes

minimum biochemical novelty and requires the fewest number of biochemical reactions.

Anammox bacteria use the energy from their catabolism for autotrophic growth. Based on the isotopic composition of cell carbon, it was proposed that anammox bacteria make use of the acetyl-coenzyme A (CoA) pathway for carbon fixation¹⁷. In the *K. stuttgartiensis* genome we detected the complete acetyl-CoA pathway (Supplementary Table 4), while all other known carbon fixation pathways remain either missing or incomplete. The acetyl-CoA pathway depends on electrons at very low redox potential (−0.41 V), usually derived from the oxidation of molecular hydrogen¹⁸. Interestingly, anammox bacteria derive their electrons from the anaerobic oxidation of nitrite to nitrate (+0.43 V) and their use of the acetyl-CoA pathway is not straightforward. To show that the acetyl-CoA pathway is really operational in anammox bacteria, we experimentally demonstrated the activity of two key enzymes (formate dehydrogenase and carbon monoxide dehydrogenase) in cell-free extracts of *K. stuttgartiensis* (Supplementary Table 5). With the genomic data we subsequently deduced a biochemical pathway that explains how the acetyl-CoA pathway can be reconciled with nitrite as electron donor for carbon fixation (Fig. 1b). In this pathway, high energy electrons from hydrazine are transferred via ferredoxin to the acetyl-CoA synthetase/CO dehydrogenase and the replenishment of the hydrazine pool to compensate for the hydrazine invested in carbon fixation requires no additional enzymes except reverse electron transport.

In the pathways of Fig. 1 there are two enzymes unique to anammox bacteria (dark blue circles and square). Hydrazine hydrolase (hh) is the enzyme which produces hydrazine from nitric

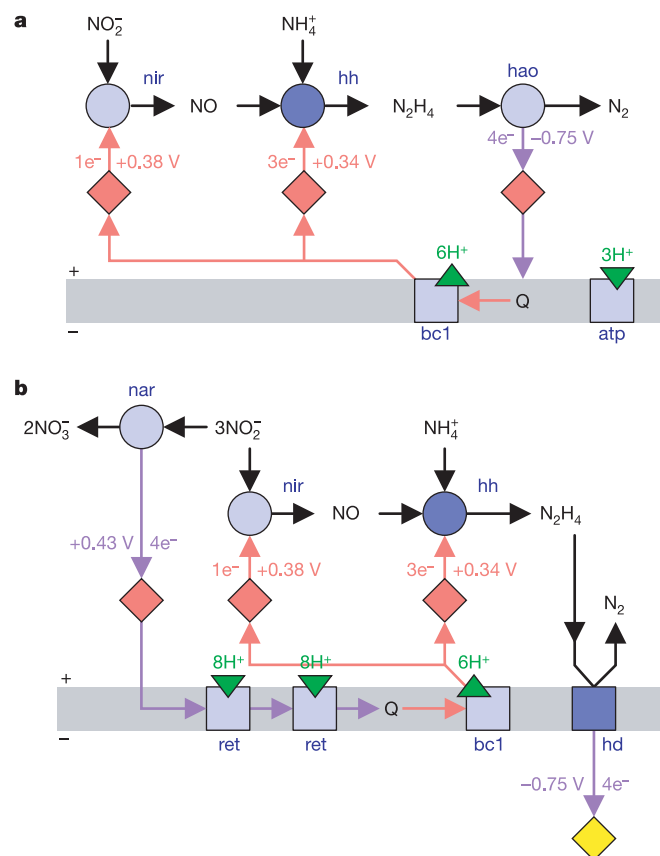


Figure 1 | Metabolic pathways of *K. stuttgartiensis* inferred from the present genomic data, previous experimental evidence and thermodynamic calculations. **a**, Anammox central catabolism with nitric oxide as intermediate, electron transport and energy conservation; **b**, combination of central catabolism with nitrate reductase to generate high potential electrons for the acetyl-CoA pathway. Three-letter abbreviations below enzymes refer to the genes and operons in Supplementary Tables 1 and 2. The dark blue ‘hh’ and ‘hd’ gene clusters are shown in Fig. 2. Red diamonds, cytochromes; yellow diamond, ferredoxin; red arrows, reductions; purple arrows, oxidations.

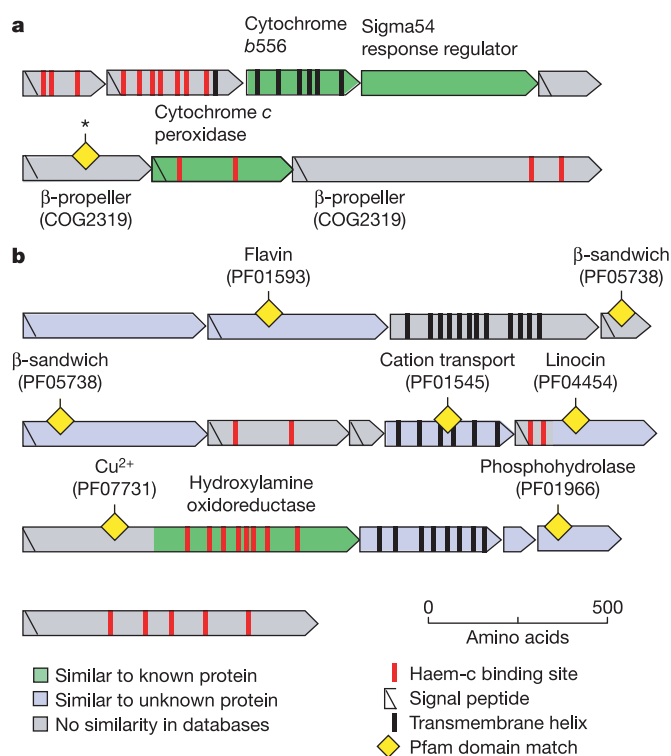


Figure 2 | The operons representing the best candidates to encode hydrazine hydrolase (hh) and hydrazine dehydrogenase (hd). **a**, Open reading frames (ORFs) *kuste2854–kuste2861* encoding a β -propeller complex (such as nitrous oxide reductase) consisting of several cytochromes. The asterisk indicates a novel cofactor binding site with weak homology to cytochrome *d* and quinoxaline binding sites. **b**, ORFs *kuste2469–kuste2483* encode a new multicopper oxidase (such as the nitrite reductase NirK), a flavin containing amine oxidase and several integral membrane proteins.

oxide and ammonium, and hydrazine dehydrogenase (hd) is the enzyme which transfers the electrons from hydrazine to ferredoxin. Interestingly, several new genes containing domains involved in electron transfer and catalysis were detected. The latter genes were mainly present in two gene clusters, which apparently code for two completely novel enzyme complexes. These are the most likely candidates to catalyse these novel steps (Fig. 2).

With a doubling time of two weeks, anammox metabolism according to Fig. 1 proceeds very slowly, and passive diffusion of protons and reactive intermediates (that is, hydrazine) over membranes is more of a problem at low enzymatic turnover. Therefore, it was proposed that the anammox bacteria have evolved ladderane lipids as the major component of their biomembranes⁵. These lipids impart unusual density and impermeability to the membrane because of their structural rigidity and size. The mode of biosynthesis is totally unknown and unprecedented because of the structural novelty and high ring strain⁶. Interestingly, in the *K. stuttgartiensis* genome fatty acid biosynthesis is represented by four gene clusters, two of which consist of a conspicuous combination of homologues of known fatty acid biosynthesis and S-adenosylmethionine (SAM) radical enzyme genes. The largest of these clusters is shown in Fig. 3. Radical SAM enzyme-encoding genes are rare in bacterial genomes and these enzymes are frequently reserved for the most difficult chemical reactions¹⁹. Although the ladderane biosynthetic pathway could not be immediately inferred, the above-mentioned gene clusters are the most promising candidates to encode this unique pathway. The encoded proteins suggest the involvement of methylation, cyclization via oxidative radical chemistry, and addition of an aromatic residue, combined with regular fatty acid elongation.

Considering the many unique adaptations, it was no surprise that previous studies based on rRNA genes indicated that anammox bacteria evolved from a common ancestor. This ancestor arose deep within the phylum Planctomycetes^{2,20}. Recently, the evolutionary position of this phylum within the tree of life has been the subject of intensive debate. The Planctomycetes are perceived as either the deepest branching bacterial phylum which forms a missing link between eukaryotes and prokaryotes^{20,21} or as a 'normal' bacterial phylum related to the intracellular parasites of the Chlamydiae²². Both

possibilities are supported by several genomic and cell-biological features^{21–25}. With the genome of *K. stuttgartiensis* and the genomic data of the planctomycete *Gemmata obscuriglobus* (<http://www.tigr.org>) it became possible to tackle this issue. We assembled a dataset of 49 concatenated amino acid sequences (Supplementary Table 6) from 90 representative bacterial genomes comprising 7,568 slowly evolving positions. A concatenated 5S–16S–23S rRNA dataset with 3,846 conserved nucleotide positions was also constructed. Phylogenetic analysis strongly supported the evolutionary grouping of Planctomycetes and Chlamydiae (Fig. 4), whereas an early divergence of the Planctomycetes from the other bacteria was not apparent, even when the trees were rooted with archaeal sequences (Supplementary Figs 4 and 5).

Next, we searched the planctomycete genomes for orthologous groups shared specifically with eukarya and other bacterial phyla. Two deduced *K. stuttgartiensis* proteins showed low similarity with eukaryal signature proteins, much less than the 17 and 10 proteins observed for *Gemmata* sp. Wa-1 and *Prostheobacter dejongei*, respectively²⁵. This finding suggests that gene transfer occurred between planctomycetes and eukaryotes after the divergence of the anammox bacteria. The planctomycetes, however, shared 17 orthologous groups with Chlamydiae, more than with any other bacterial phylum (Supplementary Table 7). Notably, all planctomycetes possess a gene encoding a protein highly similar to the well-known 60-kDa cysteine-rich outer membrane protein of Chlamydiae²³, which does not occur in any other prokaryote for which a genome sequence is available. Thus it represents a unique characteristic of the Planctomycetes–Chlamydiae superphylum. Members of this superphylum are also characterized by peptidoglycan-less cell envelopes^{21,23} even though the Chlamydiae are sensitive to penicillin and their genomes encode the complete peptidoglycan biosynthetic pathway except for penicillin-binding proteins 1 and 2 (Supplementary

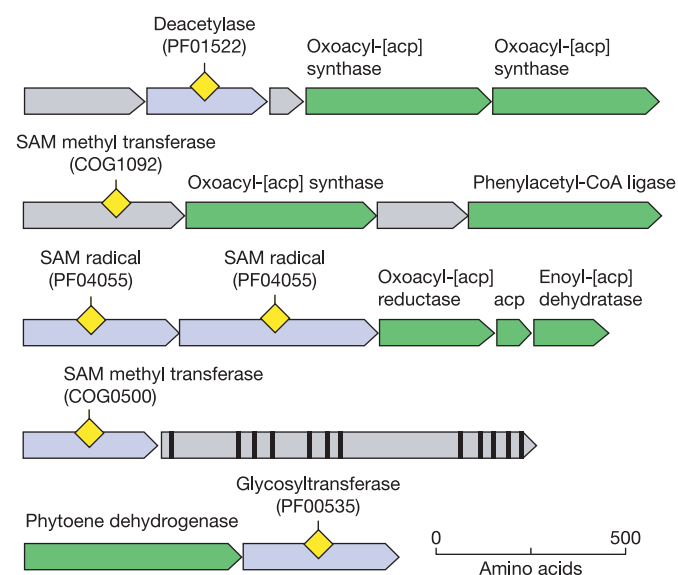


Figure 3 | The largest of four operons encoding fatty acid biosynthesis in *K. stuttgartiensis* (ORFs *kuste3352–kuste3335*). The presence of several hypothetical S-adenosylmethionine (SAM) radical and methylase proteins, highly unusual in the context of fatty acid biosynthesis, make this operon the best candidate to encode the ladderane lipid biosynthetic pathway. acp, acyl carrier protein.

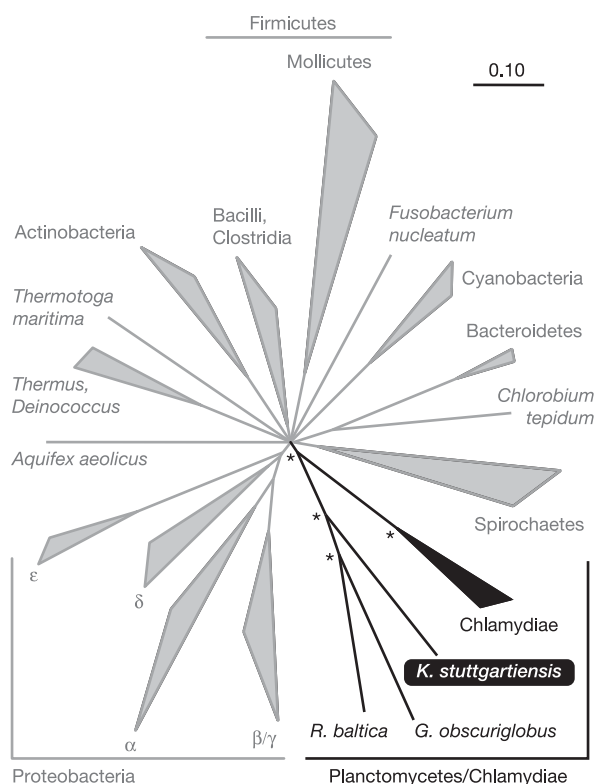


Figure 4 | Unrooted phylogenetic consensus tree based on 49 concatenated protein sequences, showing the phylogenetic positioning of *K. stuttgartiensis* within the Bacteria. Asterisks at nodes represent 100% distance- and maximum parsimony-based bootstrap support. Scale bar represents 10% estimated sequence divergence.

Table 8). In contrast to the planctomycete *Rhodopirellula baltica*, which lacks almost all peptidoglycan synthesis genes²⁴, *K. stuttgartiensis* encodes 19 out of the 21 genes required for peptidoglycan biosynthesis. This situation resembles that of the Chlamydiae, and this suggests that the common ancestor of the Planctomycetes-Chlamydiae superphylum evolved from a bacterium that was still able to synthesize a 'normal' Gram-negative cell wall.

Not so long ago, anammox bacteria were considered to be slowly growing, highly specialized chemolithoautotrophs with minor ecological importance. It has now become clear that although slow and specialized, they are global players in the biological nitrogen cycle^{1,2}. With the genome of *K. stuttgartiensis*, reconstructed from the microbial community of a laboratory bioreactor, we begin to realize that anammox bacteria could actually be generalists, not specialists; even though anammox metabolism requires distinct adaptations, it appears to be compatible with a versatile lifestyle—a lifestyle that may link the biological nitrogen cycle to the carbon and metal cycles in new ways.

METHODS

Laboratory bioreactor. The anammox bioreactor consisted of an anoxic 10-litre gas-lift reactor fed with synthetic wastewater, and was inoculated with sludge from the nitrification stage of the Dokhaven wastewater treatment plant (Rotterdam, The Netherlands). At the time of sampling (one year after inoculation), *K. stuttgartiensis* made up 73% of the microbial community^{2,26}.

DNA preparation. Bioreactor samples were immobilized in agarose plugs and DNA was extracted after incubation (3 h, 37°C) in EDTA (100 mM), NaCl (50 mM), Tris-HCl (10 mM), Na-lauroyl sarcosine (0.5%), lysozyme (1 mg ml⁻¹), mutanolysin (10 U ml⁻¹), lipase (1 mg ml⁻¹), peptidase (1 mg ml⁻¹) and β -glucuronidase (1 mg ml⁻¹) pH 8. The plugs were then incubated twice for 24 h at 50°C in EDTA (0.5 M), Na-lauroyl sarcosine (0.5%) pH 8 containing protease (500 ng ml⁻¹) and proteinase K (2 mg ml⁻¹).

Library construction. A BAC library (8,448 clones) was constructed from partial *Hind*III digests²⁷. Fosmid (6,432 clones) and shotgun-randomly-sheared-DNA-plasmid libraries were constructed with pEpiFOS5 (Epicentre) and pCDNA2 vectors respectively. Clones from all libraries were picked and bidirectionally sequenced using standard protocols.

Assembly. In a preliminary global assembly (192,713 shotgun sequence reads, performed with Phrap) a great number of small contigs were produced. After the removal of all sequences harbouring repeated motifs, a second global assembly (127,557 reads) resulted in 9,787 contigs. By BAC and fosmid bridging, 493 of these contigs were assembled into 68 supercontigs. Three genes indicative of anammox organisms (16S rDNA, a cytochrome *c* and hydroxylamine oxidoreductase) were detected on three of these supercontigs. The subsequent assignment of 17 other supercontigs to *K. stuttgartiensis* was guided by the GC content (41%) and a high percentage of BAC-end sequences (>5%, see Supplementary Methods) characteristic for these 3 supercontigs. After a second round of BAC and fosmid bridging, 5 final contigs (totalling 4,218,325 nucleotides, EMBL accession numbers CT030148, CT573071–CT573074) were finally obtained (Supplementary Fig. 2). The mean read coverage over the 5 contigs was 22 $\times$.

Confirmation of assembly. The robustness of the assembly was supported by the phylogenetic analysis of 33 conserved marker genes. In neighbour-joining or maximum parsimony trees, these markers clustered with the planctomycetes *R. baltica* or *G. obscuriglobus*. Further, 73 out of 88 shotgun sequences obtained from a 99.5% pure cell preparation of the related anammox bacterium *Brocadia anammoxidans* were most similar to sequences in the *K. stuttgartiensis* assembly when compared against a database of 180 complete bacterial genomes (including *K. stuttgartiensis*) using the program Blast.

Genome annotation. Prediction of coding sequences was performed with the programs dps/orpheus, AMIGene, glimmer/rbsfinder and genemarks/genemark.hmm. Annotation was performed in PEDANT²⁸. Genes with sequence homology (>20% identical) to genes with established functions (that is, by knockout/complementation experiments, structural analysis) were annotated as '(strongly) similar to'. Genes with sequence homology to uncharacterized genes were annotated as '(conserved) hypothetical protein'. Genes without significant homology (<20% identical) were annotated as 'unknown'. tRNA genes were identified with the tRNAscan-SE program and rRNA genes were located by homology. Genome synteny against other complete genomes was computed with LAGAN.

Phylogenetic analysis. Phylogenetic analysis of *K. stuttgartiensis* was based on concatenated data sets of protein and rRNA sequences. The data set was based on

98 representative (<97% similarity of 16S rDNA genes) genomes plus *K. stuttgartiensis* and *G. obscuriglobus*, a planctomycete currently being sequenced at The Institute for Genomic Research (TIGR). Sequences were extracted for 44 ribosomal proteins, 3 DNA-directed RNA polymerase subunits and 3 other proteins (GyrB, RecA and EF-TU) commonly used for bacterial phylogeny (Supplementary Table 6). For proteins, distance (FITCH), maximum parsimony (PROTPARS), and maximum likelihood (MOLPHY) analyses were used to construct consensus trees from concatenated alignments (CLUSTAL, default settings) with a positional conservation filter of 30%. For ribosomal RNA (5S, 16S and 23S), neighbour-joining, maximum parsimony and maximum likelihood analysis was performed in ARB after alignment with ARB-FastAligner followed by manual refinement, concatenation and application of a 50% conservation filter. Bootstrapping (100 resamplings) was performed with PHYLIP 3.61. Rooting of ribosomal protein trees with members of another domain is often not attempted owing to alignment difficulties and the risk of biases due to long-branch attraction. Here, to test whether the current data set suggests deep-branching of the Planctomycetes within the Bacteria tree, we included ribosomal protein sequences from five archaea according to the alignments of Vishwanath *et al.*²⁹. After concatenation and application of conservation filters, phylogenetic analyses were performed as described above and with PHYML³⁰ for maximum-likelihood-based phylogeny with the following evolutionary models: Dayhoff, JTT, MtRev and WAG. For each PHYML inference topology optimization was applied on a BIONJ starting tree, the proportion of invariable sites and the gamma distribution parameter were estimated, and 16 substitution rate categories were applied. These calculations were performed on Quad AMD Opteron 2.4 GHz computing nodes, equipped with 8 GB shared main memory. Methods are described and referenced in detail in the Supplementary Information.

Received 22 December 2005; accepted 15 February 2006.

- Arrigo, R. A. Marine microorganisms and global nutrient cycles. *Nature* **437**, 349–355 (2005).
- Strous, M. & Jetten, M. S. M. Anaerobic oxidation of methane and ammonium. *Annu. Rev. Microbiol.* **58**, 99–117 (2004).
- Van de Graaf, A. A., De Bruijn, P., Robertson, L. A., Jetten, M. S. M. & Kuenen, J. G. Metabolic pathway of anaerobic ammonium oxidation on the basis of N-15 studies in a fluidized bed reactor. *Microbiology (UK)* **143**, 2415–2421 (1997).
- Schalk, J., de Vries, S., Kuenen, J. G. & Jetten, M. S. M. A novel hydroxylamine oxidoreductase involved in the anammox process. *Biochemistry* **39**, 5405–5412 (2000).
- Sinninghe Damsté, J. S. *et al.* Linearly concatenated cyclobutane lipids form a dense bacterial membrane. *Nature* **419**, 708–712 (2002).
- Macitti, V. & Corey, E. J. Total synthesis of pentacyclic anammoxic acid. *J. Am. Chem. Soc.* **126**, 15664–15665 (2004).
- Lindsay, M. R. *et al.* Cell compartmentalization in planctomycetes: novel types of structural organization for the bacterial cell. *Arch. Microbiol.* **175**, 413–429 (2001).
- DeLong, E. F. Microbial community genomics in the ocean. *Nature Rev. Microbiol.* **3**, 459–469 (2005).
- Tringe, S. G. *et al.* Comparative metagenomics of microbial communities. *Science* **308**, 554–557 (2005).
- Schmid, M. C. *et al.* Molecular evidence for genus level diversity of bacteria capable of catalyzing anaerobic ammonium oxidation. *Syst. Appl. Microbiol.* **23**, 93–106 (2000).
- Von Mering, C. *et al.* STRING: known and predicted protein–protein associations, integrated and transferred across organisms. *Nucleic Acids Res.* **33**, D433–D437 (2005).
- Methe, B. A. *et al.* Genome of *Geobacter sulfurreducens*: Metal reduction in subsurface environments. *Science* **302**, 1967–1969 (2003).
- Heidelberg, J. F. *et al.* Genome sequence of the dissimilatory metal ion-reducing bacterium *Shewanella oneidensis*. *Nature Biotechnol.* **20**, 1118–1123 (2002).
- Chain, P. *et al.* Complete genome sequence of the ammonia-oxidizing bacterium and obligate chemolithoautotroph *Nitrosomonas europaea*. *J. Bacteriol.* **185**, 2759–2773 (2003).
- Brasseur, G. *et al.* Apparent redundancy of electron transfer pathways via bc1 complexes and terminal oxidases in the extremophilic chemolithoautotrophic *Acidithiobacillus ferrooxidans*. *Biochim. Biophys. Acta* **1656**, 114–126 (2004).
- Güven, D. *et al.* Propionate oxidation by and methanol inhibition of anaerobic ammonium-oxidizing bacteria. *Appl. Environ. Microbiol.* **71**, 1066–1071 (2005).
- Schouten, S. *et al.* Stable carbon isotopic fractionations associated with inorganic carbon fixation by anaerobic ammonium oxidizing bacteria. *Appl. Environ. Microbiol.* **40**, 3785–3788 (2004).
- Drake, H. L. & Daniel, S. L. Physiology of the thermophilic acetogen *Moorella thermoacetica*. *Res. Microbiol.* **155**, 422–436 (2004).
- Layer, G., Heinz, D. W., Jahn, D. & Schubert, W. D. Structure and function of radical SAM enzymes. *Curr. Opin. Chem. Biol.* **8**, 468–476 (2004).

20. Brochier, C. & Philippe, H. Phylogeny—A non hyperthermophilic ancestor for bacteria. *Nature* **417**, 244 (2002).
21. Fuerst, J. A. Intracellular compartmentation in planctomycetes. *Annu. Rev. Microbiol.* **59**, 299–328 (2005).
22. Teeling, H., Lombardot, T., Bauer, M., Ludwig, W. & Glöckner, F. O. Evaluation of the phylogenetic position of the planctomycete '*Rhodopirellula baltica*' SH 1 by means of concatenated ribosomal protein sequences, DNA-directed RNA polymerase subunit sequences and whole genome trees. *Int. J. Syst. Evol. Microbiol.* **54**, 791–801 (2004).
23. Chopra, I., Storey, C., Falla, T. J. & Pearce, J. H. Antibiotics, peptidoglycan synthesis and genomics: the chlamydial anomaly revisited. *Microbiology (UK)* **144**, 2673–2678 (1998).
24. Glöckner, F. O. *et al.* Complete genome sequence of the marine planctomycete *Pirellula* sp. Strain 1. *Proc. Natl Acad. Sci. USA* **100**, 8298–8303 (2003).
25. Staley, J. T., Bouzek, H. & Jenkins, C. Eukaryotic signature proteins of *Prostheobacter dejongei* and *Gemmata* sp Wa-1 as revealed by *in silico* analysis. *FEMS Microbiol. Lett.* **243**, 9–14 (2005).
26. Daims, H., Lückner, S. & Wagner, M. DAIME, a novel image analysis program for microbial ecology and biofilm research. *Environ. Microbiol.* **8**, 200–213 (2006).
27. Le Paslier, M. C. *et al.* Construction and characterization of a *Schistosoma mansoni* bacterial artificial chromosome library. *Genomics* **65**, 87–94 (2000).
28. Frishman, D. *et al.* Functional and structural genomics using PEDANT. *Bioinformatics* **17**, 44–57 (2001).
29. Vishwanath, P., Favaretto, P., Hartman, H., Mohr, S. C. & Smith, T. F. Ribosomal protein-sequence block structure suggests complex prokaryotic evolution with implications for the origin of eukaryotes. *Mol. Phylogenet. Evol.* **33**, 615–625 (2004).
30. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Preliminary sequence data from *G. obscuriglobus* was obtained from The Institute for Genomic Research through the website at <http://www.tigr.org>. Sequencing of *G. obscuriglobus* was accomplished with support from the United States Department of Energy. A. Cabezas is acknowledged for technical assistance. M.H., A.C. and M.W. were supported by the Austrian Ministry for Science and Education. M.T. was funded by the German Federal Ministry for Education and Research (Biolog II). H.D. was supported by the Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF). We thank B. Anneser, F. Maixner, M. Leitner, A. Pol and B. Meijerink for help in the first round annotation and M. Ott for his effort in high performance phylogenetic treeing. The Dutch anammox research was supported by the Netherlands Science Foundation for Earth and Life Science (ALW), the Foundation for Applied Research (STW) and the European Union. The Radboud University and S. Wendelaar Bonga financially supported the initial phase of the project.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.S.M.J. (m.jetten@science.ru.nl) or M.W. (wagner@microbial-ecology.net).

Energetics reveals physiologically distinct castes in a eusocial mammal

M. Scantlebury¹, J. R. Speakman², M. K. Oosthuizen¹, T. J. Roper³ & N. C. Bennett¹

Eusociality, which occurs among mammals only in two species of African mole-rat, is characterized by division of labour between morphologically distinct ‘castes’¹. In Damaraland mole-rats (*Cryptomys damarensis*), colony labour is divided between ‘infrequent worker’ and ‘frequent worker’ castes². Frequent workers are active year-round and together perform more than 95% of the total work of the colony, whereas infrequent workers typically perform less than 5% of the total work³. Anecdotal evidence suggests that infrequent workers may act as dispersers, with dispersal being limited to comparatively rare periods when the soil is softened by moisture^{4,5}. Here we show that infrequent workers and queens increase their daily energy expenditure after rainfall whereas frequent workers do not. Infrequent workers are also fatter than frequent workers. We suggest that infrequent workers constitute a physiologically distinct dispersing caste, the members of which, instead of contributing to the work of the colony and helping the queen to reproduce, build up their own body reserves in preparation for dispersal and reproduction when environmental conditions are suitable.

In communally breeding species, certain individuals forego reproduction and instead assist others to reproduce^{6,7}. This phenomenon is seen at its most extreme in eusocial organisms, where reproductive ‘queens’ are responsible for producing all the offspring of a colony. Forsaking reproduction in favour of assisting the reproduction of others is the most extreme form of altruism known among non-human animals and, as such, poses a fundamental problem for the theory of evolution^{8,9}. In most communally breeding species, the level of reproductive skew is thought to be determined by the balance between the inclusive fitness benefits arising from remaining within the natal group and the costs associated with ecological constraints on dispersal^{8,10}. This cost–benefit scenario might lead to the evolution of two alternative strategies among helpers^{11,12}. The first would involve remaining within the natal colony and expending resources to maximize indirect fitness gains; the second would involve minimizing the contribution to indirect fitness gains in the current colony, but maximizing direct fitness gains by dispersing and reproducing elsewhere whenever opportunities to disperse arise.

African mole-rats (Rodentia, Bathyergidae) include the only two known species of eusocial mammal: the naked mole-rat *Heterocephalus glaber* and the Damaraland mole-rat *Cryptomys damarensis*³. Because both species are entirely subterranean, dispersal can only be achieved by extending the burrow system, which is only possible when the soil has been softened by recent rainfall³. In Damaraland mole-rats, labour is divided between two castes: infrequent workers and frequent workers², with the infrequent workers amounting to 25–40% of the individuals within a colony³. Anecdotal evidence suggests that immediately after a period of heavy rainfall, infrequent workers make prospecting forays away from the colony in an attempt

to mate with members of other colonies⁴ or found new colonies⁵. This suggests that the primary function of infrequent workers may be to disperse when environmental conditions permit.

We tested this suggestion by measuring the daily energy expenditure of the two castes of Damaraland mole-rat worker and the queens before and after rainfall using the doubly labelled water (DLW) technique¹³. We also measured body fat content by isotope dilution¹⁴, resting metabolic rate by indirect calorimetry and sustained metabolic scope (the ratio of daily energy expenditure to resting metabolic rate¹⁵) in all individuals. We predicted that daily energy expenditure should be relatively constant for frequent workers but should increase after rainfall for infrequent workers. We also predicted that if infrequent workers were dispersers, they should have lower energy expenditure while dispersal is constrained and carry more energy reserves to fuel eventual dispersal^{4,12}. Different castes had significantly different body masses ($F_{2,71} = 23.49$, $P < 0.001$) but body mass did not differ within castes between seasons ($F_{1,71} = 0.08$, $P = 0.781$). There was no change in body mass during the DLW experimental period for any category of animal either in dry conditions (paired $t_5 = 0.96$, $P = 0.38$, $t_{13} = 1.64$, $P = 0.12$ and $t_{20} = 0.94$, $P = 0.36$ for queens, infrequent workers and frequent workers, respectively; where subscript numbers indicate degrees of freedom) or after rainfall ($t_5 = 1.35$, $P = 0.26$, $t_{12} = 1.55$, $P = 0.15$ and $t_{17} = 0.61$, $P = 0.55$ for queens, infrequent workers and frequent workers, respectively) (Table 1). Overall, queens weighed 119 ± 23.4 g ($n = 12$), infrequent workers weighed 136 ± 33.2 g ($n = 27$) and frequent workers weighed 85 ± 28.5 g ($n = 39$) (Table 1). There was no difference in body mass between male and female frequent workers (76.5 ± 21.8 g and 92.8 ± 30.8 g for males and females respectively, $F_{1,30} = 3.07$, $P = 0.09$) but male infrequent workers were significantly heavier than female infrequent workers (141.3 ± 31.9 g and 111.3 ± 26.9 g respectively, $F_{1,23} = 4.31$, $P = 0.049$). There was a significant effect of caste, but not of season or sex, on the fatness of an individual ($F_{2,58} = 3.42$, $P = 0.04$, $F_{1,58} = 1.88$, $P = 0.175$ and $F_{1,58} = 1.39$, $P = 0.244$ for caste, season and sex, respectively). Post-hoc tests (Tukey Least Significant Difference (LSD)) revealed that this was because infrequent workers were significantly fatter than frequent workers; infrequent workers carried, on average, 14.5 ± 5.3 g of fat whereas frequent workers carried only 7.5 ± 4.8 g. The difference in the amount of body fat between castes did not completely account for the difference in total body mass, suggesting that infrequent workers also had greater lean mass. Greater muscle mass may be advantageous when it comes to digging burrows, but the costs and benefits in this circumstance are complex because a larger-bodied individual may be more capable at digging, but require a larger tunnel to fit its body through.

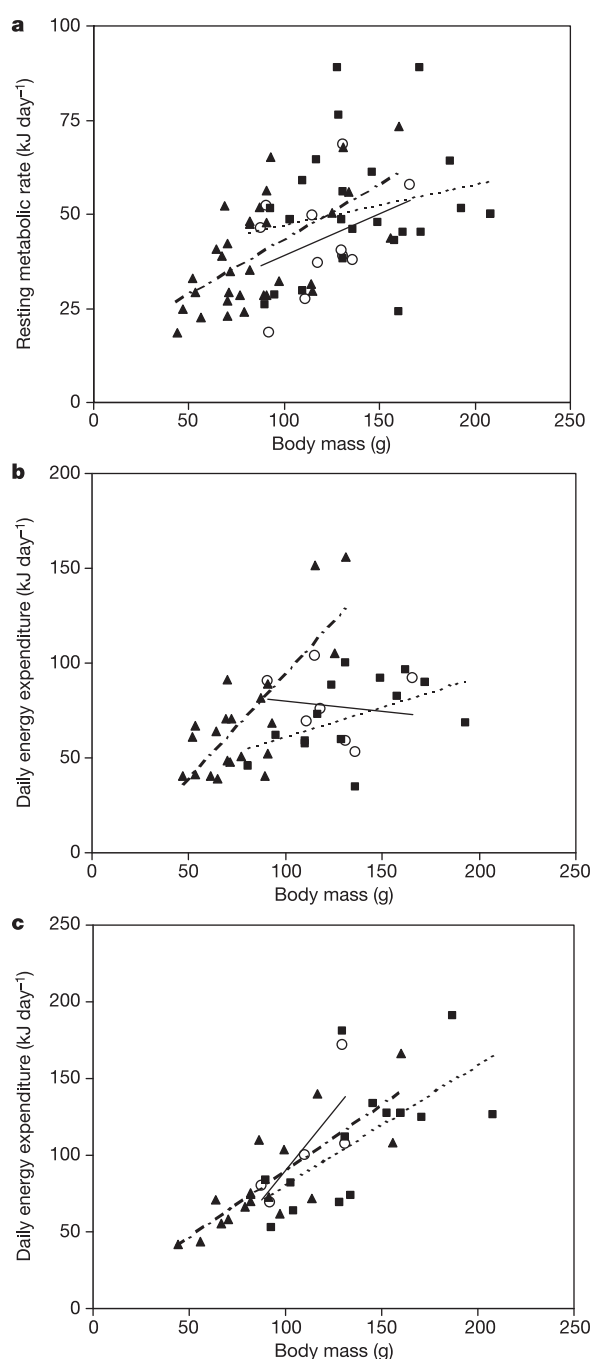
There was a significant positive relationship between resting metabolic rate and body mass ($F_{1,65} = 20.14$, $P < 0.001$; least-squares

¹Mammal Research Institute, Department of Zoology and Entomology, University of Pretoria, Pretoria 0002, South Africa. ²Aberdeen Centre for Energy Regulation and Obesity (ACERO), University of Aberdeen, School of Biological Sciences, Aberdeen AB24 2TZ, UK. ³Department of Biology & Environmental Science, University of Sussex, Brighton BN1 9QG, UK.

Table 1 | Energetic measurements

	Queen during dry conditions (mean $\pm$ s.d.)	Queen during wet conditions (mean $\pm$ s.d.)	Infrequent worker during dry conditions (mean $\pm$ s.d.)	Infrequent worker during wet conditions (mean $\pm$ s.d.)	Frequent worker during dry conditions (mean $\pm$ s.d.)	Frequent worker during wet conditions (mean $\pm$ s.d.)
<i>n</i>	6	6	14	13	21	18
Body mass (<i>x,y</i>) (g)	124, 126 $\pm$ 24, 22	110, 107 $\pm$ 23, 20	133, 132 $\pm$ 31, 31	139, 137 $\pm$ 37, 37	78, 79 $\pm$ 23, 23	93, 93 $\pm$ 32, 33
RMR (kJ day ⁻¹)	48.3 $\pm$ 12.0	48.5 $\pm$ 23.0	55.5 $\pm$ 15.2	59.1 $\pm$ 23.0	44.3 $\pm$ 15.7	43.3 $\pm$ 16.3
Mass-corrected RMR (kJ day ⁻¹)	45.4 $\pm$ 11.8	48.7 $\pm$ 20.3	49.3 $\pm$ 16.1	52.2 $\pm$ 22.3	50.8 $\pm$ 14.5	47.7 $\pm$ 12.0
DEE (kJ day ⁻¹)	77.7 $\pm$ 18.8	107.2 $\pm$ 46.0	71.9 $\pm$ 20.0	113.4 $\pm$ 42.5	70.3 $\pm$ 33.4	81.5 $\pm$ 32.5
Mass-corrected DEE (kJ day ⁻¹)	68.6 $\pm$ 24.4	105.9 $\pm$ 36.5	57.6 $\pm$ 19.1	96.1 $\pm$ 32.4	86.9 $\pm$ 24.7	90.0 $\pm$ 22.3
SusMS	1.64 $\pm$ 0.35	2.53 $\pm$ 1.23	1.39 $\pm$ 0.48	2.20 $\pm$ 1.12	1.87 $\pm$ 0.83	1.88 $\pm$ 0.34
Percentage fat	10.9 $\pm$ 2.0	8.5 $\pm$ 1.9	11.1 $\pm$ 3.5	10.3 $\pm$ 2.9	8.9 $\pm$ 3.5	7.7 $\pm$ 4.0

Mean and standard deviations (s.d.) of body mass, resting metabolic rate (RMR), mass-corrected resting metabolic rate, daily energy expenditure (DEE), mass-corrected daily energy expenditure, sustained metabolic scope (SusMS) and percentage fat for queens, infrequent workers and frequent workers. The two values of body mass (*x,y*) indicate measurements taken at the beginning and the end, respectively, of the DLW study. *n* denotes sample sizes. Resting metabolic rate was converted to kJ day⁻¹ using an oxygen-equivalent of 20.51 kJ l⁻¹ O₂; a respiratory quotient of 0.8 was assumed^{19,20}.



regression: resting metabolic rate (kJ day⁻¹) = 24.1 + 0.227 × body mass (g), $r^2 = 0.24$; Fig. 1a). However, once the effects of mass had been accounted for, there was no additional relationship between season and resting metabolic rate, or caste and resting metabolic rate ($F_{1,62} = 0.01$, $P = 0.979$ and $F_{2,62} = 0.26$, $P = 0.771$, respectively; Table 1).

There was a positive relationship between daily energy expenditure and body mass ($F_{1,75} = 41.32$, $P < 0.001$; least-squares regression: daily energy expenditure (kJ day⁻¹) = 22.5 + 0.561 × body mass (g), $r^2 = 0.36$). In addition to this overall mass effect there were complex effects of both caste and season on the levels of daily energy expenditure, which were reflected in a significant three-way interaction between season, caste and body mass on daily energy expenditure ($F_{2,65} = 3.83$, $P = 0.028$). This suggested that there was an effect of caste on the levels of daily energy expenditure, but that this effect differed with season. To explore these effects further we analysed the data for dry and wet seasons separately. In the dry season there was a positive relationship between daily energy expenditure and body mass ($F_{1,36} = 8.67$, $P = 0.006$) and a significant interaction between body mass and caste ($F_{2,36} = 6.51$, $P = 0.004$), indicating that the gradients of the changes in daily energy expenditure with mass were significantly different between castes (Fig. 1b). In absolute terms the energy demands of the different castes were 70.3 kJ day⁻¹ for the frequent workers, 71.9 kJ day⁻¹ for the infrequent workers and 77.7 kJ day⁻¹ for the queens. Once the mass effects had been removed, the mass-corrected daily energy expenditure averaged 87 kJ day⁻¹ for frequent workers, but only 58 kJ day⁻¹ and 69 kJ day⁻¹ for infrequent workers and queens, respectively, which were significantly lower than the costs for the frequent workers ($F_{2,39} = 7.09$, $P = 0.002$; Tukey LSD) (Table 1). In the wet season there was also a significant effect of mass on daily energy expenditure ($F_{1,29} = 17.58$, $P < 0.001$) but this time no significant interaction between body mass and caste ($F_{2,29} = 0.65$, $P = 0.529$); in this season the effect of caste disappeared ($F_{2,32} = 0.62$, $P = 0.546$; Tukey LSD) (Fig. 1c). The mass-corrected daily energy expenditure of the infrequent workers and queens increased during the wet season, to an average 96 kJ day⁻¹ and 106 kJ day⁻¹, respectively, so that their daily energy expenditures did not differ from that of the frequent workers, which was unchanged at an average of 90 kJ day⁻¹ (Table 1). Absolute energy demands of the three castes in the wet season were 81.5, 113.4 and 107.2 kJ day⁻¹ for frequent workers, infrequent workers and queens, respectively.

Figure 1 | Resting metabolic rate and daily energy expenditure for *C. damarensis*. **a**, Resting metabolic rate against body mass of *C. damarensis* for wet and dry periods combined. **b**, Daily energy expenditure against body mass during a dry period. **c**, Daily energy expenditure against body mass during a wet period. Open circles, solid squares and solid triangles (together with solid, dotted and dotted-dashed least-squares regression lines) denote queens, infrequent workers and frequent workers, respectively.

Because the mass-corrected resting metabolic rate was independent of caste and season, the trends in sustained metabolic scope mirrored the changes in daily energy expenditure. There was a significant interaction between season and caste for sustained metabolic scope ($F_{2,61} = 3.97$, $P = 0.025$). Post-hoc tests revealed that infrequent workers had higher sustained metabolic scope values after rainfall but frequent workers did not (Table 1).

Our results show that infrequent workers are larger than frequent workers and have a higher percentage of body fat. One possibility is that the lower energy demands of the infrequent workers occurred because they were better insulated by their larger fat stores. However, this is unlikely to explain their lower mass-corrected energy demands, because the summer temperatures in the burrow systems are relatively high¹⁶ (averaging $31.9 \pm 1.0^\circ\text{C}$). Thus, these animals routinely live above their lower critical temperatures and do not have significant thermoregulatory demands that would be alleviated by improved insulation. After a period of rainfall, infrequent workers increased their daily energy expenditure and sustained metabolic scope more than did frequent workers, suggesting that rainfall triggers energetically expensive activities such as digging^{5,17}. Taken together with previous evidence that dispersing individuals are generally larger^{4,12}, and that moist conditions stimulate reproductive activity¹⁸, these results suggest that infrequent workers constitute a dispersing caste within Damaraland mole-rat colonies. That is, during dry periods infrequent workers, instead of contributing to the work of the colony, build up stores of body fat in preparation for meeting the costs of dispersal and reproduction⁵. Because these individuals are larger and thus have greater energy demands, they place a double burden on the colony—doing virtually no work, but requiring more food.

METHODS

The study was carried out near the town of Hotazel, South Africa. We trapped mole-rats in March 2003 when conditions were dry (no rainfall for at least a month) and in March 2004 after approximately 40 mm of rain. We injected captured individuals with DLW, released them at the site of capture and then recaptured them after approximately 2–5 days. Of the 55 animals initially captured and injected during the dry period, we re-captured 41 from 8 different colonies; of the 60 individuals initially captured and injected during the wet period, we re-captured 37 from 7 colonies. On completion of experiments, animals were returned to their original capture sites. For full methodological details of measurement of daily energy expenditure and resting metabolic rate, and for details of statistical analysis, see Supplementary Information.

Received 25 August 2005; accepted 12 January 2006.

1. Jarvis, J. U. M., O'Riain, M. J., Bennett, N. C. & Sherman, P. W. Mammalian eusociality: a family affair. *Trends Ecol. Evol.* **9**, 46–51 (1994).
2. Bennett, N. C. & Jarvis, J. U. M. The social structure and reproductive biology of colonies of the mole-rat *Cryptomys damarensis* (Rodentia: Bathyergidae). *J. Mamm.* **69**, 293–302 (1988).
3. Bennett, N. C. & Faulkes, C. G. *African Mole-rats Ecology and Eusociality* 273 (Cambridge Univ. Press, Cambridge, 2000).
4. Hazell, R. W. A., Bennett, N. C., Jarvis, J. U. M. & Griffin, M. Adult dispersal in

the cooperatively breeding Damaraland mole-rat, *Cryptomys damarensis*: a case study from the Waterberg region of Namibia. *J. Zool.* **252**, 19–26 (2000).

5. Jarvis, J. U. M. & Bennett, N. C. Eusociality has evolved independently in two genera of bathyergid mole-rats—but occurs in no other subterranean mammal. *Behav. Ecol. Sociobiol.* **33**, 353–360 (1993).
6. Stacey, P. B. & Koenig, W. D. *Cooperative Breeding in Birds* (Univ. Cambridge press, Cambridge, 1990).
7. Solomon, N. G. & French, J. F. *Cooperative Breeding in Mammals* (Cambridge Univ. Press, Cambridge, 1997).
8. Emlen, S. T. in *Behavioural Ecology an Evolutionary Approach* 4th edn (eds Krebs, J. R. & Davies, N. B.) 228–253 (Blackwell, Oxford, 1997).
9. Emlen, S. T. in *Behavioural Ecology an Evolutionary Approach* 2nd edn (eds Krebs, J. R. & Davies, N. B.) 301–337 (Blackwell, Oxford, 1991).
10. Keller, L. & Reeve, H. K. Partitioning of reproduction in animal societies. *Trends Ecol. Evol.* **9**, 98–102 (1994).
11. Waterman, J. M. Mating tactics of male Cape ground squirrels, *Xerus inauris*: consequences of year-round breeding. *Anim. Behav.* **56**, 459–466 (1998).
12. O'Riain, M. J., Jarvis, J. U. M. & Faulkes, C. G. A dispersive morph in the naked mole-rat. *Nature* **380**, 619–621 (1996).
13. Speakman, J. R. *Doubly Labelled Water, Theory and Practice* (Chapman and Hall, London, 1997).
14. Scantlebury, M., Oosthuizen, M. K., Speakman, J. R., Jackson, C. R. & Bennett, N. C. Seasonal field energetics of the Hottentot golden mole (*Ambyxomys hottentotus longiceps*) at 1500m altitude. *Physiol. Behav.* **84**, 739–745 (2005).
15. Hammond, K. A. & Diamond, J. Maximal sustained energy budgets in humans and animals. *Nature* **386**, 457–462 (1997).
16. Bennett, N. C., Jarvis, J. U. M. & Davies, K. C. Daily and seasonal temperatures in the burrows of African rodent moles. *S. Afr. J. Zool.* **23**, 189–195 (1988).
17. Burland, T. M., Bennett, N. C., Jarvis, J. U. M. & Faulkes, C. G. Eusociality in African mole-rats: new insights from patterns of genetic relatedness in the Damaraland mole-rat. *Proc. R. Soc. Lond. B* **269**, 1025–1030 (2002).
18. Molteno, A. J. & Bennett, N. C. Relaxation of socially-induced reproductive inhibition in colonies of the eusocial Damaraland mole-rat: the effect of aridity as an ecological constraint promoting philopatry. *J. Zool.* **256**, 445–448 (2002).
19. Jacobs, D. S., Bennett, N. C., Jarvis, J. U. M. & Crowe, T. M. The colony structure and dominance hierarchy of the Damaraland mole-rat, *Cryptomys damarensis* (Rodentia: Bathyergidae) from Namibia. *J. Zool.* **224**, 553–576 (1991).
20. Hardy, R. N. *Temperature and Animal Life. The Institute of Biology. Studies in Biology no. 35* (Arnold, London, 1972).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This research was funded by the National Research Foundation, Republic of South Africa (N.C.B.) and a University of Pretoria Post Doctoral Research Fellowship (M.S.). We thank S. van Cutsem and H. Hjarvard de Fine Licht for assistance in the field, and F. Dalerum and G. Gutjahr for comments on the manuscript. Experimental procedures and animal husbandry practices were approved by the Animal Ethics Committee, University of Pretoria.

Author Contributions M.S., conception of research, fieldwork, organization of study site, data analysis and manuscript preparation; J.R.S., isotope analysis and contribution to manuscript preparation; M.K.O., fieldwork; T.J.R., conception of research and contribution to manuscript preparation; N.C.B., conception of research, fieldwork, organization of study site, acquisition of funding and contribution to manuscript preparation.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.S. (m.scantlebury@zoology.up.ac.za).

LETTERS

An unconventional myosin in *Drosophila* reverses the default handedness in visceral organs

Shunya Hozumi¹, Reo Maeda¹, Kiichiro Taniguchi¹, Maiko Kanai¹, Syuichi Shirakabe¹, Takeshi Sasamura¹, Pauline Spéder³, Stéphane Noselli³, Toshiro Aigaki⁴, Ryutaro Murakami⁵ & Kenji Matsuno^{1,2}

The internal organs of animals often have left–right asymmetry^{1,2}. Although the formation of the anterior–posterior and dorsal–ventral axes in *Drosophila* is well understood, left–right asymmetry has not been extensively studied. Here we find that the handedness of the embryonic gut and the adult gut and testes is reversed (not randomized) in viable and fertile homozygous *Myo31DF* mutants. *Myo31DF* encodes an unconventional myosin, *Drosophila* MyoIA (also referred to as MyoID in mammals; refs 3, 4), and is the first actin-based motor protein to be implicated in left–right patterning. We find that *Myo31DF* is required in the hindgut epithelium for normal embryonic handedness. Disruption of actin filaments in the hindgut epithelium randomizes the handedness of the embryonic gut, suggesting that *Myo31DF* function requires the actin cytoskeleton. Consistent with this, we find that *Myo31DF* colocalizes with the cytoskeleton. Overexpression of Myo61E, another myosin I (ref. 4), reverses the handedness of the embryonic gut, and its knockdown also causes a left–right patterning defect. These two unconventional myosin I proteins may have antagonistic functions in left–right patterning. We suggest that the actin cytoskeleton and myosin I proteins may be crucial for generating left–right asymmetry in invertebrates.

Mechanisms that create characteristic left–right asymmetry have been studied extensively in vertebrates². However, although the organs of many invertebrate species also have left–right asymmetry, the mechanisms by which this asymmetry arises are largely unknown. In *Drosophila*, several organs have left–right asymmetry, including the embryonic gut, the adult brain and the genitalia^{5–9}.

To identify genes involved in left–right asymmetry of the *Drosophila* embryonic gut, we performed a genetic screen using a collection of P-element lines (Gene Search, <http://gsdb.biol.metro-u.ac.jp/%7Edclust/>). The embryonic gut is composed of three major parts, the foregut, midgut and hindgut, all of which have characteristic left–right asymmetry⁵ (Fig. 1c, e, h and Table 1, row 1). We found that 75.7% of homozygous *Myo31DF^{souther}* embryos show synchronous inversion of the midgut and hindgut (Fig. 1d, f and Table 1, row 2). In these embryos, the hindgut and midgut are the mirror-image of those in wild-type embryos, rather than showing randomized patterning (binominal test, $P \leq 0.01$). In contrast, foregut handedness was normal in all cases examined, indicating that this phenotype was heterotaxial (Fig. 1i and Table 1, row 2). *Myo31DF^{souther}* is a background mutation of the Gene Search *Drosophila* line GS14508. We used deficiency mapping to map the cytological location of *Myo31DF^{souther}* to between 30D and 31F (data not shown). We then performed complementation tests between *Myo31DF^{souther}* and lines bearing mutations that map to this region. *Myo31DF^{souther}*

failed to complement *Myo31DF^{K1}* (described in ref. 10), an allele of *Myo31DF* encoding *Drosophila* MyoID^{3,4}.

The transposable element *gypsy* was inserted into the 5′-untranslated region of the *Myo31DF* gene in *Myo31DF^{souther}* (Fig. 1a). *Myo31DF^{L152}* was one of five ethylmethanesulfonate (EMS)-induced *Myo31DF* alleles isolated in a large-scale EMS mutant screen (details of the screen will be presented elsewhere). *Myo31DF^{L152}* has a base substitution that introduces a premature stop codon at amino acid 331, resulting in a putative truncated product (Fig. 1a). *Myo31DF* overexpression from *UAS-Myo31DF* driven by *byn-Gal4* in the hindgut and posterior midgut and their primordial counterparts rescued the left–right defects of *Myo31DF^{souther}* embryos (Supplementary Table 1, row 22)¹¹, indicating that *Myo31DF* was responsible for the heterotaxia. In *Myo31DF^{souther}* embryos, NP2432-driven expression of *Myo31DF* in the hindgut epithelium (Table 1, row 12), but not in other parts of the embryonic gut, such as the midgut and mesoderm (Supplementary Table 1, rows 24 and 25), was sufficient to rescue this heterotaxia, suggesting that *Myo31DF* is required in the hindgut epithelium. Furthermore, the frequency of the handedness defect was similar in *Myo31DF^{L152}* homozygous and *Myo31DF^{L152}/Df(2L)J2* embryos (Table 1, rows 3 and 4; see Methods). Thus, *Myo31DF^{L152}* is probably a null mutant of *Myo31DF*. All homozygous *Myo31DF* mutants isolated in this study were viable and fertile, with normal hindgut tissue specification, suggesting that *Myo31DF* function is largely restricted to left–right patterning (data not shown). We did not detect a maternal phenotype or *Myo31DF* gene contribution ($0.2 < P < 0.3$, χ^2 test; Supplementary Table 1, rows 5 and 6). Notably, the foregut became a mirror-image of its wild-type counterpart when *Myo31DF* was overexpressed in the entire embryo, but other parts of the gut were normal¹² (Fig. 1j and Supplementary Table 1, row 7). Therefore, we suggest that *Myo31DF* is not involved in the left–right asymmetrical development of the foregut in wild-type embryos, but can reverse foregut handedness.

We next examined the adult hindgut and testes, which are regenerated during metamorphosis. These also showed inverted handedness in the *Myo31DF* homozygote (Fig. 1l, n, p and Table 1, rows 2 and 3). In most *Myo31DF^{L152}* adults, the loop of the hindgut and spiral of the testes were reversed (binominal test, $P \leq 0.01$), although not always synchronously (Table 1, row 3). We then knocked-down the function of the *Myo31DF* gene using RNA interference (RNAi) *in vivo*. The expression of double-stranded *Myo31DF* RNA driven by *byn-Gal4* caused inversion of the adult, but not the embryonic, hindgut (Supplementary Table 1, row 9). Thus, the left–right pattern involving *Myo31DF* is not transmitted during metamorphosis.

¹Department of Biological Science and Technology, and ²Genome and Drug Research Center, Tokyo University of Science, 2641 Yamazaki, Noda, Chiba 278-8510, Japan.

³Institute of Signalling, Developmental Biology & Cancer, UMR6543-CNRS, University of Nice Sophia-Antipolis, Parc Valrose, 06108 Nice Cedex 2, France. ⁴Department of Biological Science, Tokyo Metropolitan University, 1-1 Minami-osawa, Hachioji, Tokyo 192-0397, Japan. ⁵Department of Physics, Biology, and Informatics, Yamaguchi University, 1677-1 Yoshida, Yamaguchi, Yamaguchi 753-8512, Japan.

In situ hybridization revealed *Myo31DF* expression in the amnioproctodeal invagination at stage 6 (Fig. 2a). At stages 12–14, *Myo31DF* messenger RNA was strongly detected in the primordial midgut and hindgut (Fig. 2b), and in the proventriculus, midgut and

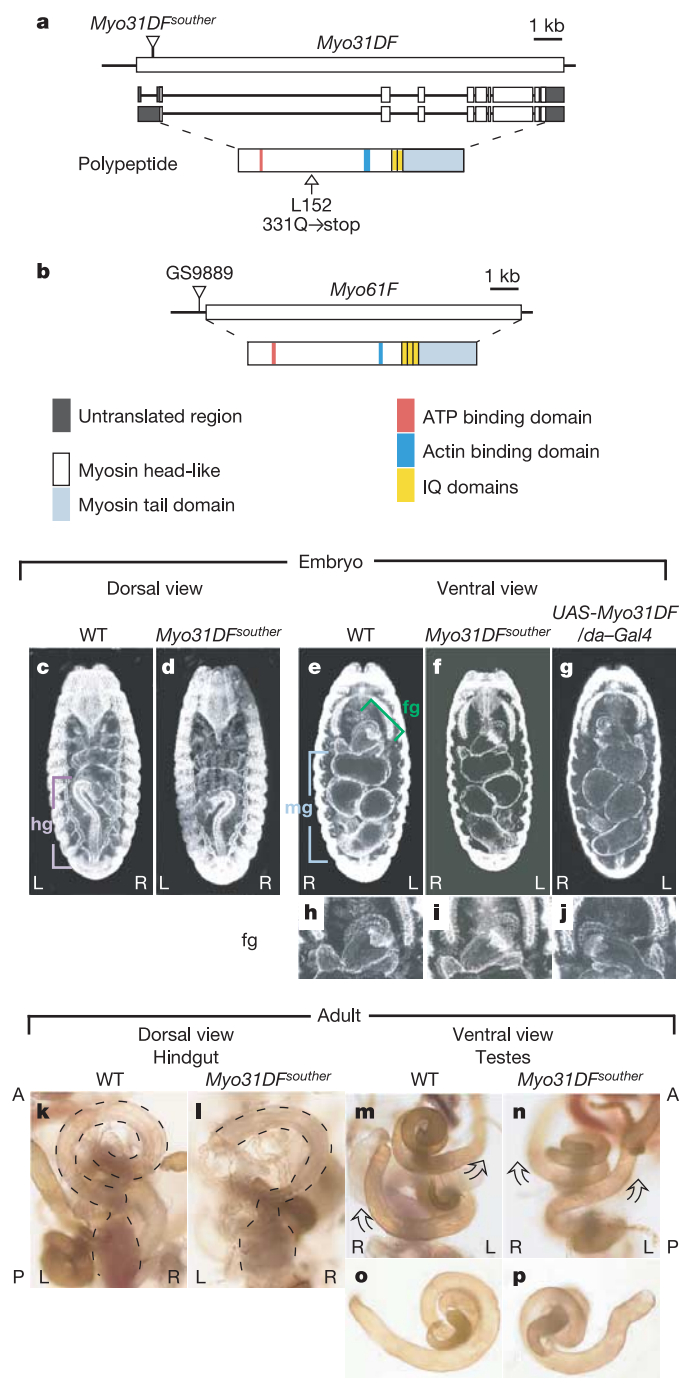


Figure 1 | *Myo31DF* mutation inverses the handedness of embryonic and adult visceral organs. **a**, Schematic of molecular lesions in *Myo31DF*^{L152} and *Myo31DF*^{souther}. **b**, In the GS9889 line, a GS vector is inserted upstream of the *Myo61F* locus. Motifs in *Myo31DF* and *Myo61F* are shown. **c**, **e**, Wild-type embryos. **d**, **f**, *Myo31DF*^{souther} embryos. **g**, Embryo overexpressing *Myo31DF* driven by *da-GAL4*. **h**–**j**, Higher-magnification views of the foregut images shown in **e**–**g**. **k**, Wild-type adult hindgut. **l**, *Myo31DF*^{souther} adult hindgut. **m**, Wild-type adult with testes located anteriorly and posteriorly, elongating to the left and right, respectively. **n**, *Myo31DF*^{souther} adult, showing mirror-image handedness of testes. **o**, **p**, Testes viewed from the ejaculatory duct in wild-type (**o**) and *Myo31DF*^{souther} (**p**) adults. Abbreviations: fg, foregut; hg, hindgut; mg, midgut; L, left; R, right; A, anterior; P, posterior.

hindgut (Fig. 2c). A sense-strand probe of *Myo31DF* gave no detectable signal (Fig. 2d). Immunostaining of wild-type embryos with an anti-Myo31DF antibody (anti-Myo31DF-1P) also labelled the midgut and hindgut (Fig. 2e–g). These signals were absent in *Myo31DF*^{souther} and *Myo31DF*^{L152} homozygotes, indicating that the staining was specific (Fig. 2h and data not shown). *Myo31DF* mRNA and protein were detected in a symmetrical pattern before the development of left–right asymmetry (data not shown). *Myo31DF* protein is expressed in the adult gut⁴. We did not detect *Myo31DF* expression in the foregut, which may account for the absence of any laterality defect in the foregut of *Myo31DF* mutants.

Myo31DF protein binds to actin in an ATP-dependent manner⁴. We next examined the co-localization of *Myo31DF* and the actin cytoskeleton in cultured *Drosophila* S2 cells. A green fluorescent protein (GFP)-tagged *Myo31DF* (*Myo31DF*–GFP) had wild-type

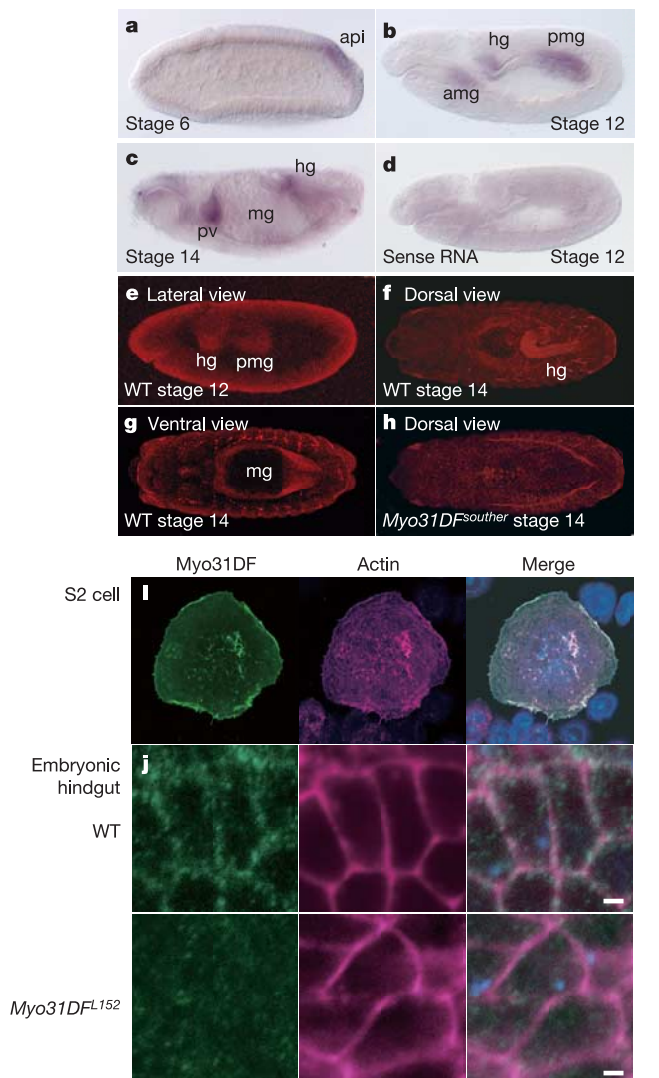


Figure 2 | Embryonic expression of *Myo31DF* and the subcellular localization of *Myo31DF*. **a**–**d**, *In situ* hybridization showing embryonic expression of *Myo31DF* at stage 6 (**a**), stage 12 (**b**) and stage 14 (**c**). **d**, Staining with a *Myo31DF* sense-strand RNA probe. **e**–**h**, *Myo31DF* detected with anti-Myo31DF-1P antibody in wild-type (WT, **e**–**g**) and *Myo31DF*^{souther} (**h**) embryos at stage 12 (**e**) and stage 14 (**f**–**h**). **i**, **j**, Localization of *Myo31DF* in S2 cells (**i**) and hindgut epithelial cells (**j**). *Myo31DF*–GFP (**i**, left) and endogenous *Myo31DF* (**j**, left) are green, actin (**i**, **j**, middle) is purple; toto3 nuclear staining is blue. Right panels show the merged images. Scale bar, 1 μ m. Abbreviations: amg, anterior midgut primordium; api, amnioproctodeal invagination; hg, hindgut; mg, midgut; pmg, posterior midgut primordium; pv, proventriculus.

Table 1 | Percentage of flies showing handedness defects

Genotype	Embryonic phenotype					Adult phenotype			
	Foregut inverse	Foregut deformed	Midgut inverse	Hindgut inverse	Midgut and hindgut deformed	Hindgut inverse	Hindgut deformed	Testis inverse	Testis deformed
Wild-type	0.5(1/200)	2.5(5/200)	0.4(1/247)	0.4(1/247)	2.4(7/247)	0.0(0/55)	0.0(0/55)	0.0(0/56)	7.1(4/56)
<i>Myo31DF^{souther}</i>	0.0(0/51)	0.0(0/51)	75.7(28/37)	75.7(28/37)	5.4(2/37)	40.7(22/54)	29.6(16/54)	73.7(28/38)	18.4(7/38)
<i>Myo31DF^{L152}</i>	1.9(1/51)	0.0(0/51)	82.0(105/128)	82.0(105/128)	6.3(8/128)	85.7(30/35)	2.9(1/35)	78.7(37/47)	19.1(9/47)
<i>Myo31DF^{L152} / Df(2L)J2</i>	-	-	57.7(30/52)	61.5(31/52)	6.0(3/52)	-	-	-	-
<i>UAS-gfp-moesin / byn-Gal4</i>	0.0(0/41)	0.0(0/41)	29.4(5/17)	47.5(28/59)	0.0(0/59)	-	-	-	-
<i>Myo31DF^{L152}; UAS-gfp-moesin / byn-Gal4</i>	0.0(0/8)	0.0(0/8)	60.0(3/5)	52.8(19/36)	0.0(0/36)	-	-	-	-
<i>UAS-Rho N19 / NP2432</i>	-	-	28.9(22/76)	21.0(16/76)	9.2(7/76)	-	-	-	-
<i>UAS-Myo61F / byn-Gal4</i>	-	-	100.0(17/17)	100.0(36/36)	0.0(0/36)	-	-	-	-
<i>UAS-Myo31DF / byn-Gal4</i>	-	-	0.0(0/47)	0.0(0/47)	0.0(0/47)	-	-	-	-
<i>UAS-dsRNA Myo61F / P{Gal4-nos.NGT}40</i>	1.5(1/65)	0.0(0/65)	7.7(5/65)*	0.0(0/65)	1.5(1/65)	-	-	-	-
Rescue experiments									
<i>(Myo31DF^{souther} background)</i>									
<i>UAS-Myo31DF / +</i>	-	-	46.3(37/80)	46.3(37/80)	2.5(2/80)	-	-	-	-
<i>UAS-Myo31DF / NP2432</i>	-	-	0.0(0/40)	0.0(0/40)	7.5(3/40)	-	-	-	-

The percentage of embryos showing defects in handedness (inverse) or deformation of organs (deformed) is shown. Genotypes of the examined embryos and adults are indicated on the left. Actual numbers of scored embryos and adults are shown in parentheses. -, values not determined. * Embryos showing partial inversion of midgut handedness are included.

function, given that its overexpression rescued *Myo31DF^{souther}* (Supplementary Table 1, row 26). *Myo31DF*–GFP co-localized with actin, mostly at cell protrusions (Fig. 2i). In epithelial cells of the hindgut, endogenous *Myo31DF* was detected as punctate staining, partly overlapping with cortical actin (Fig. 2j).

byn-Gal4-driven misexpression of GFP-tagged moesin, an actin-binding protein, in wild-type embryos caused a reduction in actin filaments in the apical region of the hindgut epithelium, where *Myo31DF* function is required¹³ (Fig. 3a,b). Notably, the midgut and hindgut always had the same handedness, but the handedness was random (not inversed), in wild-type and *Myo31DF* homozygous embryos (Table 1, rows 5 and 6). Embryo handedness was also affected by NP2432-driven GFP–moesin expressed in the hindgut epithelium only (Supplementary Table 1, row 12). However, GFP–moesin expression in the midgut only did not affect handedness (Supplementary Table 1, row 13).

To investigate further the functional link between *Myo31DF* and the actin cytoskeleton, we determined the phenocritical period for inducing left–right defects using *Myo31DF* RNAi and GFP–moesin misexpression, following the TARGET method¹⁴. Both *Myo31DF* knockdown and GFP–moesin expression in the hindgut 0–24 h after pupation caused similar defects in adult gut handedness, suggesting that *Myo31DF* and the appropriate organization of actin filaments are required at the same time (Supplementary Table 2). GFP–moesin also affected handedness in the *Myo31DF* embryo, suggesting that the default handedness, which may be manifested in the *Myo31DF* homozygote, also depends on the actin cytoskeleton (Table 1, row 6). We also examined the involvement of three Rho GTPase family proteins, Rho, Rac and Cdc42, which regulate the organization of the actin cytoskeleton¹⁵. Expression of dominant-negative forms of these proteins, especially Rho, in the hindgut induced synchronous left–right defects in the embryonic midgut and hindgut (Table 1, row 7 and Supplementary Table 1, rows 14 and 16). Together, our results suggest that *Myo31DF* depends on the actin cytoskeleton to generate left–right asymmetry.

Cell division and cell death do not occur during left–right asymmetric development of the hindgut¹⁶. We therefore speculated that the rearrangement of hindgut epithelial cells may be part of this

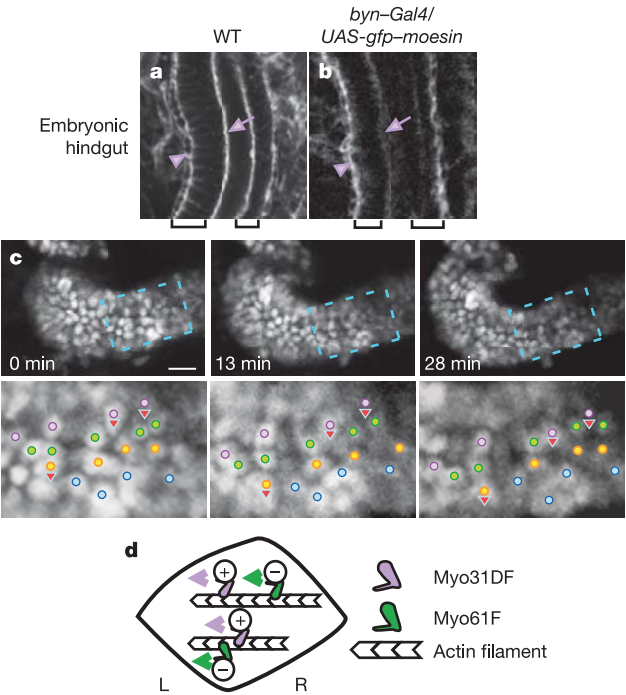


Figure 3 | *Myo31DF* is dependent on the actin cytoskeleton to develop left–right asymmetry. **a, b**, Actin localization in hindgut epithelium of wild-type (**a**) and GFP–moesin-overexpressing (**b**) embryos. The basal surface is indicated with an arrowhead, the apical surface with an arrow, and hindgut epithelial cell layers are indicated with brackets. **c**, Time-lapse analysis of hindgut epithelial cells. Dots of each colour indicate a single row of cells at *t* = 0 min. Images obtained at 0, 13 and 28 min are shown. The lower panels show a higher magnification image of the boxed area in each upper panel. Note that each row of cells tilts (with right becoming upper), which coincides with the left-handed rotation of the hindgut. Cells showing intercalation are indicated with arrowheads showing the direction of their movement. Scale bar, 10 μ m. **d**, Model for myosin I protein functions in *Drosophila* left–right asymmetry. *Myo31DF* and *Myo61F* transport left–right determinants with opposite activities.

process. To test this possibility, we performed a time-lapse analysis. The position of each cell was visualized by labelling the nucleus with GFP. Cell rearrangement, which coincided with the left–right bias associated with left-handed rotation of the hindgut, was suggested by the significant intercalation of some cells (arrowheads in Fig. 3c).

Another myosin I protein, Myo61F (*Drosophila* MyoIB, also referred to as MyoIC in mammals), has been reported in *Drosophila*^{3,4,17}. Myo61F protein is detected in the embryo and adult gut⁴. To test whether *Myo61F* is also involved in left–right asymmetry, we overexpressed *Myo61F* using *UAS-Myo61F* or GS9889 driven by *byn-Gal4* (Gene Search; Fig. 1b). Unexpectedly, Myo61F overexpression resulted in inversion of the midgut and hindgut in both cases (Table 1, row 8, and data not shown). In contrast, *Myo31DF* overexpression did not affect the handedness of these organs (Table 1, row 9). These results suggest that Myo31DF and Myo61F have antagonistic functions in creating the left–right asymmetry of these organs. The involvement of Myo61F in left–right asymmetry is also supported by our finding that its knockdown by RNAi results in the left–right defect in the embryonic midgut (Table 1, row 10).

We have found that homozygous *Myo31DF* embryos show reversed handedness of embryonic and adult visceral organs, which may represent the default state of left–right asymmetry in *Drosophila*. This situation is similar to the function of the *sinistral* gene in the freshwater snail, *Limnea* (although the *sinistral* gene is required maternally)^{1,18}. Normal handedness is still seen in 25% of *Myo31DF* homozygotes. We speculate that some other myosin gene(s) has a redundant function in left–right patterning. Inversion of the anteroposterior axis does not affect laterality, suggesting that left–right patterning occurs zygotically⁷; this is consistent with the zygotic function of *Myo31DF*. Our results also suggest that an actin-based mechanism, which can align itself to either an anteroposterior–dorsoventral reference or the pre-existing sinistral handedness, exists to direct the rotation of the hindgut epithelium. As myosin I proteins are involved in vesicular transport¹⁹, we propose that Myo31DF and Myo61F, which on the basis of their structures are believed to move to the plus ends of actin filaments, carry left–right determinants with opposite activities (Fig. 3d). Thus, both left–right determinants would be concentrated in the plus ends of actin filaments that have a hypothetical planar polarity. In the *Myo31DF* mutant, only the opposing determinant is concentrated here, which reverses the handedness. According to our model, disruption in actin organization would result in left–right randomization, as we indeed observed experimentally.

METHODS

***Drosophila* stocks.** We used Canton-S as the wild-type *Drosophila* strain. *Myo31DF^{south}* and *Myo31DF^{L152}* are newly characterized *Myo31DF* mutations. GS9889 is a Gene Search line. *Df(2L)J2* has a deletion between 31B and 32A (Bloomington Stock Center). The following GAL4 drivers were used: *byn-Gal4* drives GAL4 expression in the hindgut and the posterior midgut primordium at stage 8, and in the longitudinal visceral mesoderm at stage 11 (ref. 11). *da-Gal4* drives uniform GAL4 expression¹². *how^{24B}* drives GAL4 expression in the mesoderm primordium from stage 11 (ref. 20). *48Y* drives GAL4 expression in the anterior and posterior midgut primordium from stage 10 (ref. 21). NP2432 drives GAL4 expression in the hindgut epithelium from stage 9 (fly stock from National Institute of Genetics, <http://www.shigen.nig.ac.jp/fly/nigfly>; data not shown). NP5021 drives GAL4 expression in the whole gut from stage 9 (fly stock from National Institute of Genetics; data not shown). *P{Gal4-nos.NGT}40* expresses *Gal4* mRNA maternally²².

The following *UAS* lines were used: *UAS-gfp-moesin*, which encodes a fusion protein of the actin-binding domain of moesin and GFP (provided by S. Hayashi)¹³; *UAS-Rho N19* (ref. 23); *UAS-Rac1 N17* (ref. 24); *UAS-Cdc42 F89* (ref. 25); *UAS-Myo31DF* (provided by S. Noselli); *UAS-Myo31DF-GFP* and *UAS-2 × Myo31DF^{RNAi}*. The last two *UAS* lines are described in the accompanying paper¹⁰.

UAS-Myo61F lines carry a pUAST transformation vector that has an insertion of the entire open reading frame of *Myo61F* cDNA (clone GH04201). *UAS-dsRNA Myo61F* lines carry a construct in which *Myo61F* cDNA corresponding to nucleotides 446–848 of GH04201 was inserted into pUAST as an inverted repeat

with an interruption of the *Myo61F* fourth intron. *UAS-dsRNA Myo61F* was maternally driven by *P{Gal4-nos.NGT}40*. The genotypes of each embryo were determined using appropriate blue-balancers, such as *CyO*, *P{en1}wg^{en1}* and *TM3, fzh-lacZ*. All crosses, except those used for the TARGET analysis, were performed at 25 °C on standard *Drosophila* medium.

Analysis of phenocritical periods. We used the TARGET method to determine the phenocritical periods for inducing the left–right defect of the adult gut by knocking down *Myo31DF* and expressing of GFP–moesin¹⁴. We used a temperature-sensitive GAL80, a suppressor of GAL4, to control the activity of GAL4 driven by *byn-Gal4*. This allowed us to express a double-stranded RNA corresponding to a portion of *Myo31DF* mRNA and to express GFP–moesin in a temporally specific manner. Flies were cultured at 18 °C, and pupae were collected at 24-h intervals. Pupae collected 0–24, 24–48 and 48–72 hours after pupation were cultured at 30 °C until eclosion.

Histological analyses of embryos. Antibody staining of *Drosophila* embryos was performed as previously described²⁶. Embryos were photographed using Zeiss Axioskop2 plus or Zeiss Pascal microscopes. Primary antibodies were a mouse anti-Fasciclin III antibody (7G10, Developmental Studies Hybridoma Bank; 1:100 dilution)²⁷ and a rabbit anti-Myo31DF-1P antibody (1:10 dilution). The anti-Myo31DF-1P rabbit serum was raised against a glutathione S-transferase fusion protein containing Myo31DF amino acids 728–990 (also described in ref. 10). The anti-Myo31DF-1P antibody was affinity purified using the same Myo31DF polypeptide with a His-tag. Secondary antibodies were Cy3 anti-rabbit IgG, Cy3 anti-mouse IgG, and anti-rabbit Alexa488 (Jackson ImmunoResearch; used at 1:1,000 dilution). TOTO3 was used as a nuclear marker (Molecular Probes; 1:200 dilution). Actin filaments were stained with rhodamine–phalloidin²⁸ (Molecular Probes; 1:200 dilution). Whole-mount *in situ* hybridization was carried out as described²⁹. A digoxigenin-labelled RNA probe was prepared from a full-length cDNA template of *Myo31DF* using a DIG RNA labelling kit (Roche).

Microscopic analysis of embryonic gut handedness. Genotypes of the embryos that were selected for study were identified by the lack of blue-balancers after β-galactosidase staining²⁶. The handedness of the foregut, midgut and hindgut were scored at stages 15–16, stage 16, and stages 14–16, respectively.

Cell culture and staining. *Drosophila* S2 cells were plated and cultured on concanavalin A (Sigma) as previously described³⁰. The cells were co-transfected with pUAS-*Myo31DF-GFP* and pAyGAL4 using CellFECTIN (Invitrogen). S2 cells were fixed and stained as described³⁰. The secondary antibody used was anti-rabbit Alexa488 (Jackson ImmunoResearch; 1:200 dilution). Staining with rhodamine–phalloidin (1:40 dilution) was carried out as described²⁸. The nuclear marker was toto3 (1:200 dilution).

Time-lapse analysis. Live embryos expressing NP5021-driven *UAS-GFP^{nls}* in the whole gut were mounted with FL-100 (Shin-Etsu Chemical Co.). Images were collected on a Zeiss Pascal microscope at intervals of 25 s from stage 13 onwards, and were processed with Zeiss LSM Image Browser and Adobe Photoshop software.

Received 20 September 2005; accepted 2 February 2006.

- Wood, W. B. Left–right asymmetry in animal development. *Annu. Rev. Cell Dev. Biol.* **13**, 53–82 (1997).
- Mercola, M. & Levin, M. Left–right asymmetry determination in vertebrates. *Annu. Rev. Cell Dev. Biol.* **17**, 779–805 (2001).
- Gillespie, P. G. et al. Myosin-I nomenclature. *J. Cell Biol.* **155**, 703–704 (2001).
- Morgan, N. S., Heintzelman, M. B. & Mooseker, M. S. Characterization of Myosin-IA and Myosin-IB, two unconventional myosins associated with the *Drosophila* brush border cytoskeleton. *Dev. Biol.* **172**, 51–71 (1995).
- Hayashi, T. & Murakami, R. Left–right asymmetry in *Drosophila melanogaster* gut development. *Dev. Growth Differ.* **43**, 239–246 (2001).
- Ligoxygakis, P., Strigini, M. & Averof, M. Specification of left–right asymmetry in the embryonic gut of *Drosophila*. *Development* **128**, 1171–1174 (2001).
- Hayashi, M. et al. Left–right asymmetry in the alimentary canal of the *Drosophila* embryo. *Dev. Growth Differ.* **47**, 457–460 (2005).
- Pascual, A., Huang, K., Neven, J. & Pr  at, T. Neuroanatomy: Brain asymmetry and long-term memory. *Nature* **427**, 605–606 (2004).
- Adam, G., Perrimon, N. & Noselli, S. The retinoic-like juvenile hormone controls the looping of left–right asymmetric organs in *Drosophila*. *Development* **130**, 2397–2406 (2003).
- Sp  der, P.,       , G. & Noselli, S. Type ID unconventional myosin controls left–right asymmetry in *Drosophila*. *Nature* doi:10.1038/nature04623 (this issue).
- Iwaki, D. D. & Lengyel, J. A. A Delta–Notch signaling border regulated by Engrailed/Inverted repression specifies boundary cells in the *Drosophila* hindgut. *Mech. Dev.* **114**, 71–84 (2002).
- Wodarz, A., Hinz, U., Engelbert, M. & Knust, E. Expression of crumbs confers apical character on plasma membrane domains of ectodermal epithelia of *Drosophila*. *Cell* **82**, 67–76 (1995).

13. Edwards, K. A., Demsky, M., Montague, R. A., Weymouth, N. & Kiehart, D. P. GFP-moesin illuminates actin cytoskeleton dynamics in living tissue and demonstrates cell shape change during morphogenesis in *Drosophila*. *Dev. Biol.* **191**, 103–117 (1997).
14. McGuire, S. E. *et al.* Spatiotemporal rescue of memory dysfunction in *Drosophila*. *Science* **302**, 1765–1768 (2003).
15. Etienne-Manneville, S. & Hall, A. Rho GTPases in cell biology. *Nature* **420**, 629–635 (2002).
16. Iwaki, D. D., Johansen, K. A., Singer, J. B. & Lengyel, J. A. *drumstick*, *bowl*, and *lines* are required for patterning and cell rearrangement in the *Drosophila* embryonic hindgut. *Dev. Biol.* **240**, 611–626 (2001).
17. Gillespie, P. G., Wagner, M. C. & Hudspeth, A. J. Identification of a 120 kd hair-bundle myosin located near stereociliary tips. *Neuron* **11**, 581–594 (1993).
18. Shibasaki, Y., Shimizu, M. & Kuroda, R. Body handedness is directed by genetically determined cytoskeletal dynamics in the early embryo. *Curr. Biol.* **14**, 1462–1467 (2004).
19. Huber, L. A. *et al.* Both calmodulin and unconventional myosin Myr4 regulate membrane trafficking along the recycling pathway of MDCK cells. *Traffic* **1**, 494–503 (2000).
20. Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
21. Martin-Bermudo, M. D., Dunin-Borkowski, O. M. & Brown, N. B. Specificity of PS integrin function during embryogenesis resides in the α subunit extracellular domain. *EMBO J.* **16**, 4184–4193 (1997).
22. Tracey, W. D. Jr, Ning, X., Klingler, M., Kramer, S. G. & Gergen, J. P. Quantitative analysis of gene function in the *Drosophila* embryo. *Genetics* **154**, 273–284 (2000).
23. Strutt, D. I., Weber, U. & Mlodzik, M. The role of RhoA in tissue polarity and Frizzled signaling. *Nature* **387**, 292–295 (1997).
24. Luo, L., Liao, Y. J., Jan, L. Y. & Jan, Y. N. Distinct morphogenetic functions of similar small GTPases: *Drosophila* Drac1 is involved in axonal outgrowth and myoblast fusion. *Genes Dev.* **8**, 1787–1802 (1994).
25. Eaton, S., Auvinen, P., Luo, L., Jan, Y. N. & Simons, K. CDC42 and Rac1 control different actin-dependent processes in the *Drosophila* wing epithelium. *J. Cell Biol.* **131**, 151–164 (1995).
26. Sullivan, W., Ashburner, M. & Hawley, R. S. *Drosophila Protocols* (Cold Spring Harbor Laboratory Press, New York, 2000).
27. Patel, N. H., Snow, P. M. & Goodman, C. S. Characterization and cloning of fasciclin III: a glycoprotein expressed on a subset of neurons and axon pathways in *Drosophila*. *Cell* **48**, 975–988 (1987).
28. Frydman, H. M. & Spradling, A. C. The receptor-like tyrosine phosphatase Lar is required for epithelial planar polarity and for axis determination within *Drosophila* ovarian follicles. *Development* **128**, 3209–3220 (2001).
29. Jiang, J., Kosman, D., Ip, Y. T. & Levine, M. The dorsal morphogen gradient regulates the mesoderm determinant twist in early *Drosophila* embryos. *Genes Dev.* **5**, 1881–1891 (1991).
30. Rogers, S. L., Rogers, G. C., Sharp, D. J. & Vale, R. D. *Drosophila* EB1 is important for proper assembly, dynamics, and positioning of the mitotic spindle. *J. Cell Biol.* **158**, 873–884 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Developmental Studies Hybridoma Bank at the University of Iowa, the Bloomington Stock Center, and the *Drosophila* Genetic Resource Center at the Kyoto Institute of Technology. We thank J. Lengyel for *byn-Gal4* and S. Hayashi for *UAS-gfp-moesin* transgenic strains. This work was supported by grants-in-aid from the Japanese Ministry of Education, Culture, Sports and Science.

Author Contributions Experimental work was performed by S.H., R.M., K.T., M.K., S.S., T.S., P.S. and T.A. Data analysis was by S.H., R.M., K.T. and K.M., and project planning was coordinated by S.N., R.M. and K.M.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to K.M. (matsuno@rs.noda.tus.ac.jp).

Type ID unconventional myosin controls left-right asymmetry in *Drosophila*

Pauline Spéder^{1*}, Géza Ádám^{1,2*} & Stéphane Noselli¹

Breaking left–right symmetry in Bilateria embryos is a major event in body plan organization that leads to polarized adult morphology, directional organ looping, and heart and brain function^{1–4}. However, the molecular nature of the determinant(s) responsible for the invariant orientation of the left–right axis (situs choice) remains largely unknown. Mutations producing a complete reversal of left–right asymmetry (situs inversus) are instrumental for identifying mechanisms controlling handedness, yet only one such mutation has been found in mice (*inversin*)⁵ and snails^{6,7}. Here we identify the conserved type ID unconventional myosin 31DF gene (*Myo31DF*) as a unique situs inversus locus in *Drosophila*. *Myo31DF* mutations reverse the dextral looping of genitalia, a prominent left–right marker in adult flies. Genetic mosaic analysis pinpoints the A8 segment of the genital disc as a left–right organizer and reveals an anterior–posterior compartmentalization of *Myo31DF* function that directs dextral development and represses a sinistral default state. As expected of a determinant, *Myo31DF* has a trigger-like function and is expressed symmetrically in the organizer, and its symmetrical overexpression does not impair left–right asymmetry. Thus *Myo31DF* is a dextral gene with actin-based motor activity controlling situs choice. Like mouse *inversin*⁸, *Myo31DF* interacts and colocalizes with β -catenin, suggesting that situs inversus genes can direct left–right development through the adherens junction.

In wild-type males, the genital plate, to which the spermiduct is attached, undergoes a 360° clockwise (dextral) rotation when viewed from the posterior pole^{9,10} (Fig. 1a). This directional looping is reminiscent of other coiling processes such as mammalian heart tube looping and snail spiral development. As in other species, one direction is dominant among the Drosophilidae, the dextral rotation, with no sinistral species reported to date¹⁰ (Supplementary Table 1). Here we have identified an insertional mutation, *KG02246*, which shows a striking inverted phenotype when combined with deficiencies covering the 31DF genomic region. In *KG02246/Df(2L)Exel7048* or *KG02246/Df(2L)J3* males, genitalia rotation is variable, with ~60% of individuals showing sinistral rotation (Table 1). Imprecise excision of *KG02246* generated two genomic deletions, *KG02246*¹ and *KG02246*² (Fig. 1f), both of which presented a stronger, 100% inverted (sinistral) genitalia rotation phenotype (Fig. 1b–e and Table 1). Thus, *KG02246* alleles identify the first situs inversus mutations in *Drosophila* and provide genetic evidence for the existence of a left–right axis in this organism.

Deficiency mapping (Table 1) and molecular cloning (Fig. 1f) revealed that *KG02246* mutations target the *Myo31DF* gene (also known as *CG7438*), which encodes *Drosophila* myosin IA (ref. 11). In support of this finding, we found that inducible RNA interference (RNAi) against *Myo31DF* in the genital disc mimicked *KG02246* mutations. Furthermore, expression of wild-type *Myo31DF* was

sufficient to fully rescue the *KG02246*¹ rotation phenotype (Table 1; see below). Accordingly, we renamed the situs inversus mutations *Myo31DF*^K (for *KG02246*), *Myo31DF*^{K1} (for *KG02246*¹) and *Myo31DF*^{K2} (for *KG02246*²).

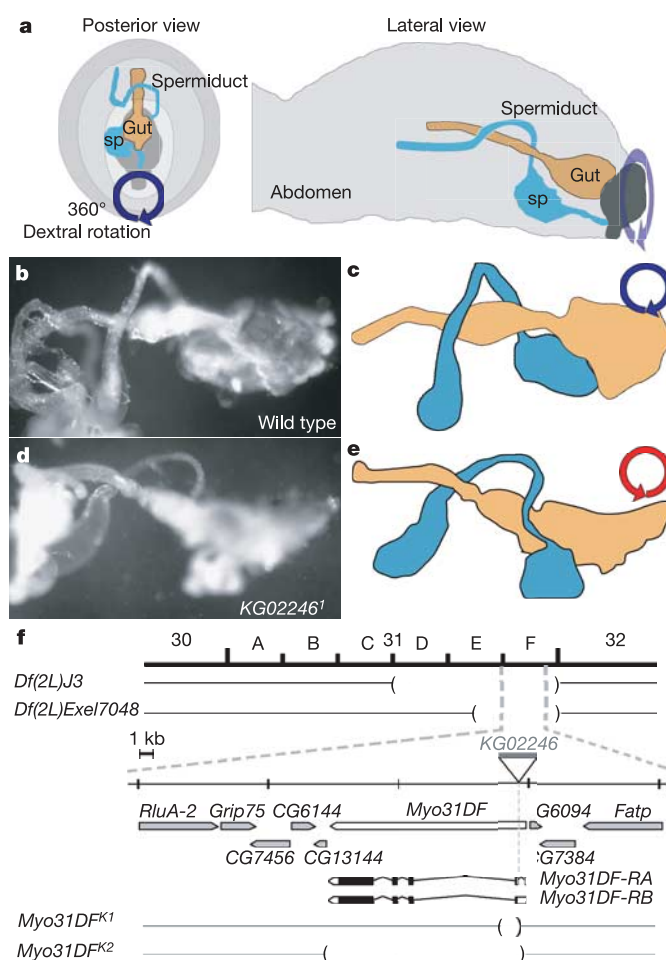


Figure 1 | Myosin31DF controls directional organ looping in *Drosophila*. **a**, Schematic posterior and lateral views of an adult male abdomen, showing the coiling of the spermiduct (blue) around the gut (orange) as a result of a 360° dextral rotation of the genital plate. sp, sperm pump. **b–e**, Dissected wild-type (**b**) and *Myo31DF*^{K1} (**d**) abdomens, showing the reversal of spermiduct looping in *Myo31DF*^{K1} flies. Panels **c** and **e** show schematic views of **b** and **d**, respectively. Colour code as in **a**. **f**, The genomic region 31E–F on chromosome 2 containing the *Myo31DF* gene.

¹Institute of Signalling, Developmental Biology & Cancer, UMR6543-CNRS, University of Nice Sophia-Antipolis, Parc Valrose, 06108 Nice Cedex 2, France. ²Institute of Genetics, Biological Research Center of the Hungarian Academy of Sciences, PO Box 521, Szeged H-6701, Hungary.

*These authors contributed equally to this work.

Drosophila Myo31DF is a conserved myosin belonging to the Myo1D family. It is known to interact with F-actin¹¹ and it colocalizes with actin-rich structures in different tissues (Supplementary Fig. 1a–g). After mouse inversin, *Drosophila* Myo31DF represents the second situs inversus gene to be molecularly identified. The mouse inversin gene (*Invs*) encodes a protein with ankyrin repeats and two IQ domains^{12,13} that bind calmodulin, a Ca^{2+} -dependent regulatory protein¹⁴. Like inversin, Myo31DF contains two IQ domains essential for its function. Indeed, a Myo31DF form lacking IQ domains (Myo31DF^{ΔIQ}) was unable to rescue Myo31DF mutations (Table 1).

To determine the function of Myo31DF in genitalia rotation, we examined its expression in the genital disc, the precursor of adult genitalia. The genital disc is composed of segments A8, A9 and A10, each with an anterior and a posterior compartment^{15–18} (Fig. 2a). Immunostaining of wild-type flies with two polyclonal antibodies directed against overlapping regions of the Myo31DF tail domain, anti-Myo31DF-1P (Fig. 2) and anti-Myo31DF-3P (Supplementary Figs 2 and 3), or using a Myo31DF–Gal4 enhancer-trap line (NP1548; Supplementary Fig. 3) revealed symmetrical expression of Myo31DF in a double chevron-like pattern restricted to the ventral domain of

the male genital disc. This expression, starting in third instar larvae and remaining unchanged during this stage (Supplementary Fig. 3), was absent in Myo31DF^{K2} mutant discs (Fig. 2c). Double staining with an A8-specific marker in the genital disc (*tsh-Gal4 > myr-RFP*) indicated that Myo31DF is expressed in the A8 segment (Fig. 2d), with one chevron in the posterior compartment (Fig. 2e, f) and the other in the anterior compartment. Consistently, expression of two copies of an inhibitory RNA gene ($2 \times \text{Myo31DF}^{\text{RNAi}}$) specifically in the A8 segment led to loss of Myo31DF expression (Fig. 2g) and to inverted phenotypes that mimicked Myo31DF mutations (Fig. 3a and Table 1). Together, these data identify the A8 segment as a left–right organizer that is required for situs choice in *Drosophila*.

To investigate the relationship between the anterior–posterior and left–right axes, we mapped the putative domain(s) of Myo31DF function by selectively silencing the gene in different compartments, using specific Gal4 lines driving $2 \times \text{Myo31DF}^{\text{RNAi}}$ (Fig. 3a, Table 1 and Supplementary Fig. 4). The combinatorial removal of Myo31DF function in the anterior and/or posterior domains (summarized in Fig. 3b) led to distinct phenotypes, indicating a dual function for the Myo31DF protein in A8. First, blocking Myo31DF function posteriorly in A8 (using *hh-Gal4* or *en-Gal4*) resulted in a striking

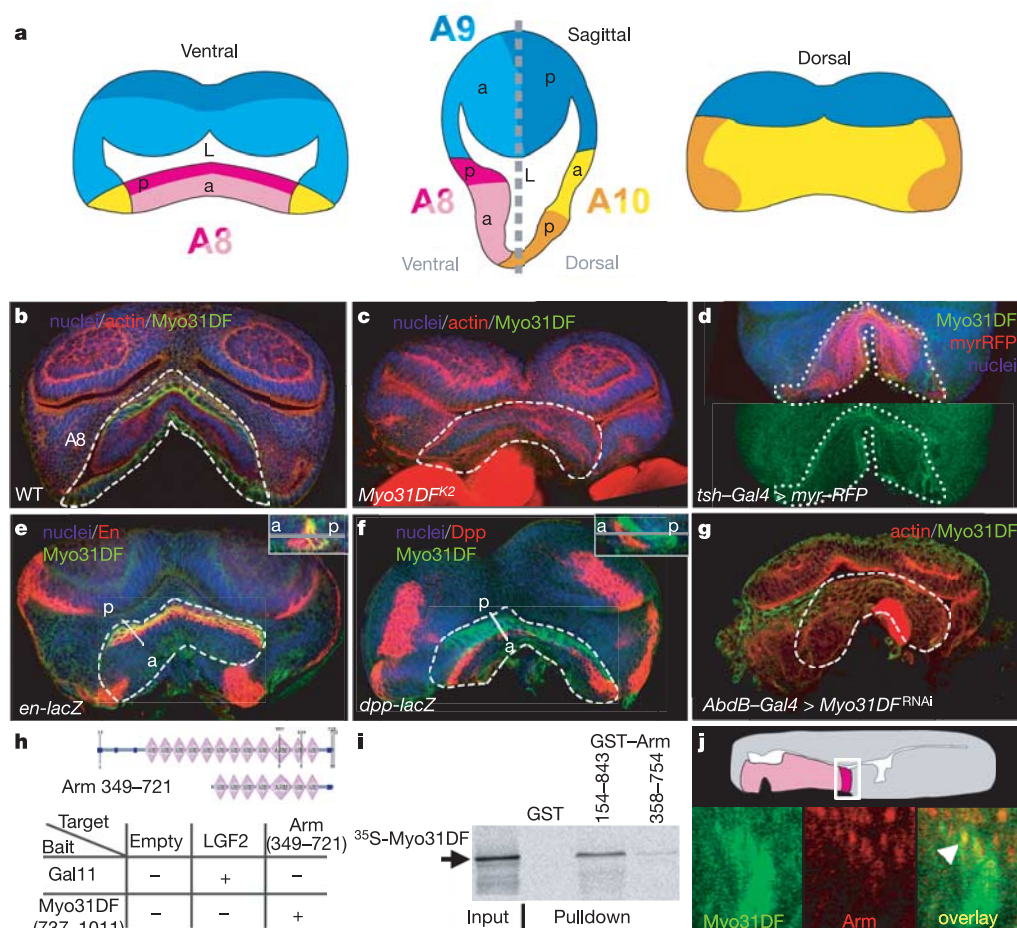


Figure 2 | Spatial expression of Myo31DF in the A8 segment of the genital disc and its interaction with Armadillo. **a**, Segmental organization of the genital disc showing A8, A9 and A10 segments with their anterior (a, light colours) and posterior (p, dark colours) compartments. L, lumen. **b–g**, Ventral views, with the A8 segment outlined (dotted line). Myo31DF expression was detected using anti-Myo31DF-1P antibodies. **b**, Endogenous Myo31DF is expressed in a double chevron-like pattern in the ventral region of the wild-type (WT) disc. **c**, Myo31DF expression is lost in the Myo31DF^{K2} genital disc. **d**, Colocalization of Myo31DF with the A8-specific marker *tsh*. **e, f**, Myo31DF expression is compartmentalized in A8. *en-lacZ* and *dpp-lacZ* are posterior and anterior markers, respectively. Bar indicates the position of the z section framed in the inset. Grey line in the inset is the x, y plane shown in **e** and **f**. **g**, Targeted expression of $2 \times \text{Myo31DF}^{\text{RNAi}}$ in A8 abolishes Myo31DF expression. **h**, Two-hybrid interactions identifying Armadillo (Arm, residues 349–721) as a binding partner of the Myo31DF tail domain. Gal11 and LGF2 were used as controls. **i**, GST-pulldown experiments showing the physical interaction between Myo31DF and Armadillo. **j**, Sagittal confocal section through the A8 posterior compartment (framed region in the schematic genital disc) showing colocalization of Myo31DF (green) and Armadillo (red).

Dpp-lacZ are posterior and anterior markers, respectively. Bar indicates the position of the z section framed in the inset. Grey line in the inset is the x, y plane shown in **e** and **f**. **g**, Targeted expression of $2 \times \text{Myo31DF}^{\text{RNAi}}$ in A8 abolishes Myo31DF expression. **h**, Two-hybrid interactions identifying Armadillo (Arm, residues 349–721) as a binding partner of the Myo31DF tail domain. Gal11 and LGF2 were used as controls. **i**, GST-pulldown experiments showing the physical interaction between Myo31DF and Armadillo. **j**, Sagittal confocal section through the A8 posterior compartment (framed region in the schematic genital disc) showing colocalization of Myo31DF (green) and Armadillo (red).

non-rotated genitalia phenotype (Table 1 and Supplementary Fig. 5). This finding was confirmed in a complementary experiment using *dpp-Gal4* to rescue *Myo31DF^{K1}* in the anterior compartment (Table 1 and Fig. 3a). The absence of situs choice observed in these experiments indicates that posterior *Myo31DF* has an instructive role in dextral looping. Second, blocking *Myo31DF* function solely in the anterior compartment (in *dpp-Gal4* > 2 × *Myo31DF^{RNAi}* or in *Myo31DF^{K1}/Myo31DF^{K1}*; *hh-Gal4* > *Myo31DF* rescued males; Table 1 and Fig. 3a) led to partial dextral rotation, suggesting a permissive role for *Myo31DF* in this compartment. Comparing this outcome to the effect of concomitant removal of both anterior and posterior functions—the only context that led to complete reversal—indicates that the function of *Myo31DF* in the anterior compartment is to repress sinistral looping (compare rows 3 and 4 in Fig. 3b). These experiments demonstrate that *Myo31DF* is essential both anteriorly and posteriorly. A model illustrating the dual function of *Myo31DF* in regulating dextral development within the A8 segment is presented in Fig. 3c. In this model, A8 contains information to specify both sinistral and dextral rotation, with sinistral information being anterior and dextral information posterior. Dextral information is dominant over sinistral information, and *Myo31DF* function is required both posteriorly, to induce dextral development, and anteriorly, to repress sinistral development. Only in the absence of *Myo31DF* can sinistral development occur as a default state.

As expected of a left–right determinant with a function that precedes asymmetry, *Myo31DF* is expressed symmetrically in A8 (Fig. 2). In addition, as with mouse *inversin*¹⁹, symmetrical overexpression does not lead to left–right defects (*AbdB-Gal4*, *tsh-Gal4* and *ptc-Gal4* lines; Table 1 and data not shown). Another predicted feature of left–right determinants is their temporally restricted, trigger-like function, which is later relayed by mechanisms acting to maintain the initial symmetry-breaking event. To test this, we carried out temperature-shift experiments using a genetically engineered temperature-sensitive *Myo31DF* allele (see Methods). Single temperature-shift experiments with 24-h (Fig. 3d) or 6-h (Fig. 3e) resolution indicated that *Myo31DF* function is required at day 6 of development, between 126 and 132 h. Double temperature-shift experiments allowed to determine that *Myo31DF* function is

required for as little as 3 h within this period (Fig. 3f). These data provide high temporal resolution, allowing the temporal mapping of a left–right symmetry-breaking event and demonstrating that situs choice depends on a peak of *Myo31DF* function in the left–right organizer.

Cilia have emerged as important cellular structures for generating left–right asymmetry in vertebrate embryos. To address their possible contribution to invertebrate left–right determination, we stained genital discs with GT335, an antibody that labels glutamylated tubulin in cilia across species, including *Drosophila*²⁰. GT335 did not detect any cilia in genital discs (data not shown). Additionally, mutations in the conserved *Rfx* gene, which controls the formation of ciliated neurons in *Drosophila*²¹ and left–right asymmetry in mouse²², did not affect genitalia rotation (data not shown). Together, these data suggest that cilia are not involved in dextral development in *Drosophila*. We propose that *Drosophila* uses primarily the actin cytoskeleton to determine left–right asymmetry. Consistently, the small GTPase *Drac1* and the JNK pathway, known regulators of the actin cytoskeleton, showed specific genetic interactions with *Myo31DF* (ref. 23 and Supplementary Fig. 1h). Notably, actin is also important for situs choice in snails²⁴, suggesting that in invertebrates the actin cytoskeleton has a central role in left–right determination.

To start investigating how *Myo31DF* might act to determine left–right asymmetry, we used two-hybrid screening to identify *Myo31DF* interactors or cargo(es). Using the *Myo31DF* tail domain (amino acids 737–1011) as bait (see Methods), we found several positive clones encoding a carboxy-terminal fragment containing ARM repeats 6–12 of the Armadillo/β-catenin protein (Fig. 2h). This interaction was direct, as shown by glutathione S-transferase (GST)-pulldown experiments using purified proteins or S2 cell extracts (Fig. 2i and Supplementary Fig. 6a). Furthermore, endogenous *Myo31DF* and a functional *Myo31DF*–GFP (green fluorescent protein) fusion protein (Table 1 and Supplementary Fig. 6b) colocalized with Armadillo at the adherens junctions in A8 (Fig. 2j), suggesting that the two proteins can interact *in vivo*. Notably, *inversin* has been shown to colocalize and interact with β-catenin in vertebrate epithelial cells⁸, indicating that an interaction with β-catenin is a

Table 1 | Summary of rotation phenotypes

	Dextral		No rotation 0°	Sinistral	
	Complete (+360°)	Partial (<+360°)		Partial (<−360°)	Complete (−360°)
Wild type	100	0	0	0	0
<i>KG02246/Df(2L)J3</i>	17	3	18	58	4
<i>KG02246/Df(2L)Exel7048</i>	0	5	40	47	9
<i>Myo31DF^{K1}</i>	0	0	0	42	58
<i>Myo31DF^{K2}</i>	0	0	0	23	77
<i>UAS-2 × Myo31DF^{RNAi}</i> crossed to*					
<i>elav-Gal4</i>	100	0	0	0	0
<i>cad-Gal4</i>	100	0	0	0	0
<i>dpp-Gal4</i>	53	32	8	7	0
<i>en-Gal4</i>	10	4	86	0	0
<i>hh-Gal4</i>	0	0	100	0	0
<i>AbdB-Gal4</i>	0	0	30	69	1
<i>ptc-Gal4</i>	0	0	4	57	39
<i>tsh-Gal4</i>	0	0	0	55	45
Rescue experiments (<i>Myo31DF^{K1}</i> background)					
<i>ptc-Gal4</i> > <i>Myo31DF[†]</i>	100	0	0	0	0
<i>ptc-Gal4</i> > <i>Myo31DF-GFP</i>	100	0	0	0	0
<i>AbdB-Gal4</i> > <i>Myo31DF</i>	70	27	2	1	0
<i>AbdB-Gal4</i> > <i>Myo31DF-GFP</i>	95	4	1	0	0
<i>dpp-Gal4</i> > <i>Myo31DF</i>	0	10	82	8	0
<i>hh-Gal4</i> > <i>Myo31DF</i>	84	16	0	0	0
<i>ptc-Gal4</i> > <i>Myo31DF^{ΔIQ}‡</i>	0	0	0	89	11

Numbers represent the percentage of male flies with the indicated genitalia rotation phenotype, from wild type (360° dextral rotation) to completely inverted (360° sinistral rotation). Unless otherwise noted, all crosses were performed at 25 °C.

*Crosses performed at 30 °C.

‡Cross performed at 20 °C.

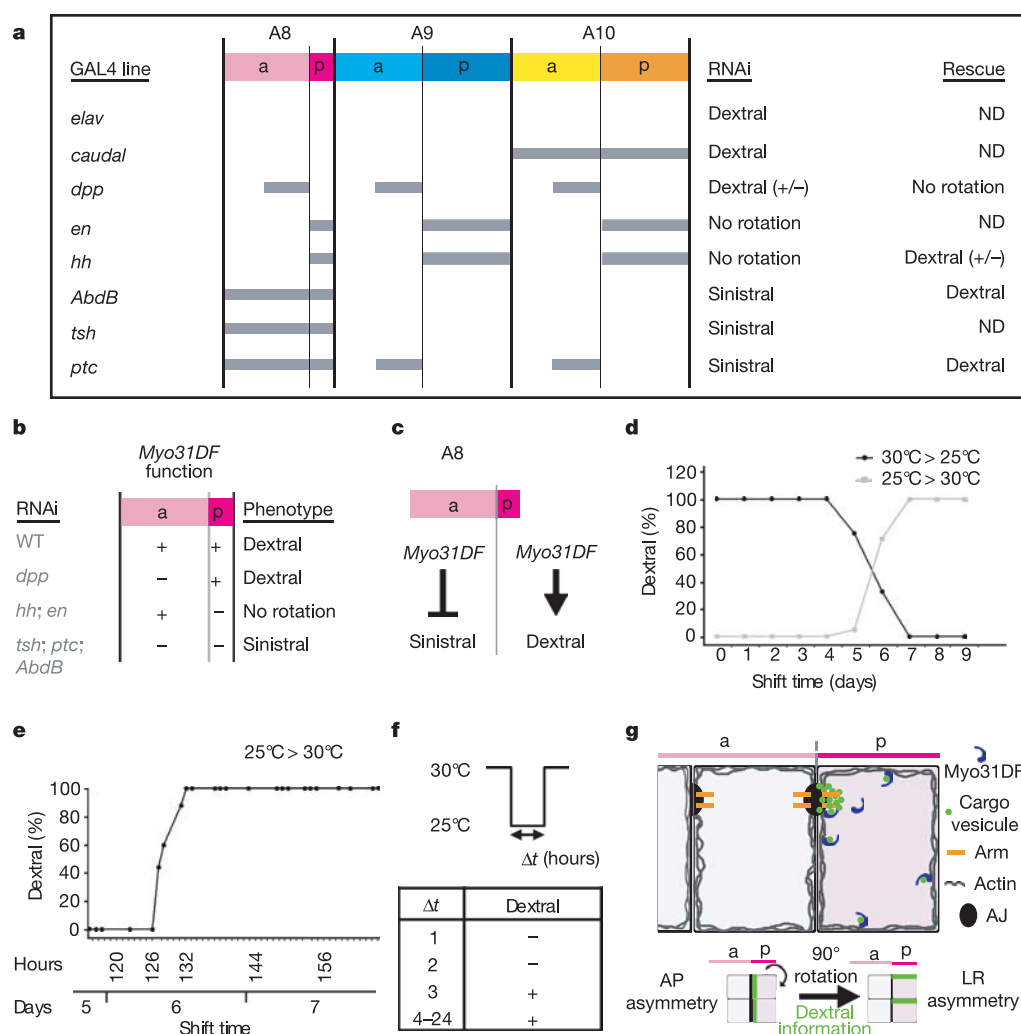


Figure 3 | Spatial and temporal organization of *Myo31DF* function in the left-right organizer. **a**, GAL4 expression domains (grey bars) and their effect on genitalia rotation upon crossing with $2 \times Myo31DF^{RNAi}$ ('RNAi') or in rescue experiments using *UAS-Myo31DF* ('Rescue') (see Table 1 for details). ND, not determined. **b**, Summary of compartmental silencing of *Myo31DF* in A8. **c**, Model of compartmental *Myo31DF* function. **d–f**, Single (**d**, **e**) or double (**f**) temperature-shift experiments in *tub-Gal80^{ts}*;

AbdB-Gal4, *UAS-Myo31DF^{RNAi}* male flies. **d**, Permissive (25 °C) to restrictive (30 °C) (grey) and restrictive to permissive (black) temperature-shifts. **e**, Requirement for *Myo31DF* function between 126–132 h in development. **f**, Double temperature-shift experiments. **g**, Proposed model of *Myo31DF* function at the adherens junction (AJ) (see text for details). AP, anterior-posterior; LR, left-right.

common feature of both known situs inversus genes. These results are consistent with a demonstrated role of N-cadherin in left-right asymmetry²⁵.

How could anterior-posterior organization of *Myo31DF* and interaction with Armadillo account for left-right asymmetry? The actin cable network might serve as a track for *Myo31DF* to deliver specific cargoes or vesicles at the adherens junction. Indeed, we found that *Myo31DF* interacts with dynamin (Supplementary Fig. 1i), consistent with the role of rat *MyoID* in vesicular transport²⁶. The anterior-posterior boundary itself creates an asymmetric junction that could serve as a scaffold for *Myo31DF* to assemble a dextral-specific complex on the anterior side of posterior *Myo31DF*-expressing cells (green in Fig. 3g). Anterior-posterior asymmetry of dextral information could later be translated into left-right asymmetry through the remodelling of cell contacts (cell intercalation or rotation), as seen in other epithelia²⁷. For example, 90° rotation of epithelial cells relative to the main body axis is observed in the *Drosophila* eye imaginal disc²⁸. A similar, 90° planar rotation of *Myo31DF*-expressing cells would result in left-right orientation of the dextral junctional complex in the tissue (Fig. 3g). In this working model, a

short pulse of *Myo31DF* activity would be essential for spatially restricting the dextral junction, as we have observed.

METHODS

Genetics. A description of genetic markers and chromosomes can be found at FlyBase (<http://flybase.bio.indiana.edu>). *AbdB-Gal4* was a gift from E. Sanchez-Herrero. *Myo31DF^{K1}* and *Myo31DF^{K2}* alleles were made by excision of the KG02246 P-element using standard protocols.

DNA cloning. Constructs were made as follows. For pUAS-*Myo31DF*, the full-length *Myo31DF* cDNA (SD01662) was cut with *EcoRI* and *XhoI* and ligated into the *EcoRI* and *XhoI* sites of the transformation vector pUAS. For pUAS-*Myo31DF-GFP*, the full-length coding region of *Myo31DF* was cut with *EcoRI* and *NotI* and ligated into the *EcoRI* and *NotI* sites of the transformation vector pUAS-GFP. For pUAS-*Myo31DF^{RNAi}*, the *BglIII-XhoI* fragment from SD01662 (nucleotides 3211–3780) was cloned into the *BamHI-XhoI* sites of the transformation vector SympUAS. pUAS-*Myo31DF^{ΔIQ}* was made by cloning the *EcoRI-XhoI* fragment from SD01662 with deleted nucleotides 2829–2952 (residues 695–736). For each construct, several independent transgenic lines were generated and tested.

Antibodies, immunostaining and imaging. We raised antibodies against two different GST fusion proteins containing *Myo31DF* tail residues 728–990

(Myo31DF-1P antibody) or 823–1011 (Myo31DF-3P antibody). The sera were purified on His-tagged peptide columns and tested by western blotting to ensure specificity (Supplementary Fig. 2). We carried out immunostaining according to standard procedures, using the following antibodies and markers: purified rabbit polyclonal antibodies against Myo31DF (1:10 to 1:100 dilutions), anti-Armadillo monoclonal antibody (DSHB; 1:20 dilution); Alexa546- and Alexa488-conjugated secondary antibodies (Molecular Probes; 1:500 dilution), phalloidin-TRITC (Sigma; 0.2 $\mu\text{g ml}^{-1}$) and Hoechst33258 (Sigma; 3.3 $\mu\text{g ml}^{-1}$).

Temperature-shift experiments. Conditional expression of Myo31DF was achieved using Gal80^{ts} in *tub-Gal80^{ts}*; *AbdB-Gal4*; *UAS-2 × Myo31DF^{RNAi}* males. At 25 °C (permissive for Gal80 and Myo31DF functions), males had a wild-type phenotype, but at 30 °C (restrictive for Gal80 and Myo31DF functions) all males had non-rotated genitalia. In single temperature-shift experiments, 1-h staged first instar larvae were grown at a given temperature and shifted once to the permissive or restrictive temperature after 24 h (Fig. 3d) or 6 h (Fig. 3e), and the phenotype of adult males was determined. In double temperature-shift experiments, 24-h staged embryos were grown at 30 °C and shifted to 25 °C for a limited period of time (1 h–24 h; Δt in Fig. 3f) before being shifted back to 30 °C until hatching. The presence of wild-type males indicated that the duration of the shift was sufficient to restore Myo31DF function.

Two-hybrid and GST-pulldown. Two-hybrid screening was performed using the Bacteriomatch II system (240065, Stratagene) according to the manufacturer's instructions. We used the Myo31DF tail domain as bait (residues 737–1011) and a *Drosophila* embryo cDNA library (982600, Stratagene) as the target. Equivalent amounts of GST, GST-Arm (residues 154–843) and GST-Arm (residues 358–754) were used to pull down *in vitro* translated, ³⁵S-labelled Myo31DF protein (TnT, Promega). After binding for 3 h at 4 °C, beads were washed four times with 1 ml lysis buffer (150 mM NaCl, 50 mM Tris pH 7.5, 0.2% Triton and protease inhibitor cocktail) and resuspended in Laemmli buffer. Samples were run on an SDS-PAGE gel and revealed using Phosphorimager.

Imaging. Images were taken on an LSM510 META confocal microscope (Zeiss) using $\times 25$ 0.80 NA and $\times 40$ 1.3 NA oil immersion objectives, and assembled using Adobe Photoshop 7.0.

Received 21 September 2005; accepted 1 February 2006.

1. Bisgrove, B. W., Morelli, S. H. & Yost, H. J. Genetics of human laterality disorders: insights from vertebrate model systems. *Annu. Rev. Genomics Hum. Genet.* **4**, 1–32 (2003).
2. Levin, M. Left–right asymmetry in embryonic development: a comprehensive review. *Mech. Dev.* **122**, 3–25 (2005).
3. Hamada, H., Meno, C., Watanabe, D. & Saijoh, Y. Establishment of vertebrate left–right asymmetry. *Nature Rev. Genet.* **3**, 103–113 (2002).
4. Raya, A. & Belmonte, J. C. Sequential transfer of left–right information during vertebrate embryo development. *Curr. Opin. Genet. Dev.* **14**, 575–581 (2004).
5. Morgan, D. *et al.* Inversin, a novel gene in the vertebrate left–right axis pathway, is partially deleted in the *inv* mouse. *Nature Genet.* **20**, 149–156 (1998).
6. Ueshima, R. & Asami, T. Evolution: single-gene speciation by left–right reversal. *Nature* **425**, 679 (2003).
7. Freeman, G. & Lundelius, J. The developmental genetics of dextrality and sinistrality in the gastropod *Lymnaea peregra*. *Wilhelm Roux's Archives* **191**, 69–83 (1982).
8. Nurnberger, J., Bacallao, R. L. & Phillips, C. L. Inversin forms a complex with catenins and N-cadherin in polarized epithelial cells. *Mol. Biol. Cell* **13**, 3096–3106 (2002).
9. Adam, G., Perrimon, N. & Noselli, S. The retinoic-like juvenile hormone controls the looping of left–right asymmetric organs in *Drosophila*. *Development* **130**, 2397–2406 (2003).
10. Gleichauf, R. Anatomie und variabilität des geschlechtsapparates von *Drosophila melanogaster* (Meigen). *Z. Wiss. Zool.* **148**, 1–66 (1936).
11. Morgan, N. S., Heintzelman, M. B. & Mooseker, M. S. Characterization of myosin-1A and myosin-1B, two unconventional myosins associated with the *Drosophila* brush border cytoskeleton. *Dev. Biol.* **172**, 51–71 (1995).
12. Morgan, D. *et al.* The left–right determinant inversin has highly conserved ankyrin repeat and IQ domains and interacts with calmodulin. *Hum. Genet.* **110**, 377–384 (2002).
13. Eley, L. *et al.* A perspective on inversin. *Cell Biol. Int.* **28**, 119–124 (2004).
14. Bahler, M. & Rhoads, A. Calmodulin signaling via the IQ motif. *FEBS Lett.* **513**, 107–113 (2002).
15. Casares, F., Sanchez, L., Guerrero, I. & Sanchez-Herrero, E. The genital disc of *Drosophila melanogaster*. 1. Segmental and compartmental organization. *Dev. Genes Evol.* **207**, 216–228 (1997).
16. Chen, E. H. & Baker, B. S. Compartmental organization of the *Drosophila* genital imaginal discs. *Development* **124**, 205–218 (1997).
17. Freeland, D. E. & Kuhn, D. T. Expression patterns of developmental genes reveal segment and parasegment organization of *D. melanogaster* genital discs. *Mech. Dev.* **56**, 61–72 (1996).
18. Sanchez, L., Casares, F., Gorfinkel, N. & Guerrero, I. The genital disc of *Drosophila melanogaster*. 2. Role of the genes *hedgehog*, *decapentaplegic* and *wingless*. *Dev. Genes Evol.* **207**, 229–241 (1997).
19. Watanabe, D. *et al.* The left–right determinant inversin is a component of node monocilia and other 9+0 cilia. *Development* **130**, 1725–1734 (2003).
20. Bobinnec, Y., Marcaillou, C. & Debec, A. Microtubule polyglutamylation in *Drosophila melanogaster* brain and testis. *Eur. J. Cell Biol.* **78**, 671–674 (1999).
21. Dubruille, R. *et al.* *Drosophila* regulatory factor X is necessary for ciliated sensory neuron differentiation. *Development* **129**, 5487–5498 (2002).
22. Bonnafant, E. *et al.* The transcription factor RFX3 directs nodal cilium development and left–right asymmetry specification. *Mol. Cell. Biol.* **24**, 4417–4427 (2004).
23. Macias, A. *et al.* PVF1/PVR signaling and apoptosis promotes the rotation and dorsal closure of the *Drosophila* male terminalia. *Int. J. Dev. Biol.* **48**, 1087–1094 (2004).
24. Shibasaki, Y., Shimizu, M. & Kuroda, R. Body handedness is directed by genetically determined cytoskeletal dynamics in the early embryo. *Curr. Biol.* **14**, 1462–1467 (2004).
25. Garcia-Castro, M. I., Vielmetter, E. & Bronner-Fraser, M. N-Cadherin, a cell adhesion molecule involved in establishment of embryonic left–right asymmetry. *Science* **288**, 1047–1051 (2000).
26. Huber, L. A. *et al.* Both calmodulin and the unconventional myosin Myr4 regulate membrane trafficking along the recycling pathway of MDCK cells. *Traffic* **1**, 494–503 (2000).
27. Bertet, C., Sulak, L. & Lecuit, T. Myosin-dependent junction remodelling controls planar cell intercalation and axis elongation. *Nature* **429**, 667–671 (2004).
28. Strutt, H. & Strutt, D. EGF signaling and ommatidial rotation in the *Drosophila* eye. *Curr. Biol.* **13**, 1451–1457 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the Bloomington Stock Center, B. Durand, B. Edde, A. Laurençon, V. van de Bor, J. P. Vincent and M. L. Cariou for materials and fly lines; E. Sanchez-Herrero for *AbdB-Gal4*; Y. Bellaiche and M. Morgan for GST-Arm protein; C. Mionnet for help with GST-pulldown assays; and C. Featherstone, P. Follette, E. Sanchez-Herrero, M. Suzanne, L. Wolpert and laboratory members for critically reading the manuscript. This work was supported by the Centre National de la Recherche Scientifique (CNRS), the Ministère de l'éducation et de la Recherche (ACI), the Hungarian National Scientific Research Fund (OTKA), the Association pour la Recherche contre le Cancer (ARC) and the EMBO Young Investigator Programme.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.N. (noselli@unice.fr).

LETTERS

Toll-dependent selection of microbial antigens for presentation by dendritic cells

J. Magarian Blander^{1†} & Ruslan Medzhitov¹

Dendritic cells constitutively sample the tissue microenvironment and phagocytose both microbial and host apoptotic cells^{1–4}. This leads to the induction of immunity against invading pathogens or tolerance to peripheral self antigens, respectively^{5–9}. The outcome of antigen presentation by dendritic cells depends on their activation status, such that Toll-like receptor (TLR)-induced dendritic cell activation makes them immunogenic, whereas steady-state presentation of self antigens leads to tolerance^{5,6,8,10}. TLR-inducible expression of co-stimulatory signals is one of the mechanisms of self/non-self discrimination^{5,11}. However, it is unclear whether or how the inducible expression of co-stimulatory signals would distinguish between self antigens and microbial antigens when both are encountered by dendritic cells during infection^{6,8}. Here we describe a new mechanism of antigen selection in dendritic cells for presentation by major histocompatibility complex class II molecules (MHC II) that is based on the origin of the antigen. We show that the efficiency of presenting antigens from phagocytosed cargo is dependent on the presence of TLR ligands within the cargo. Furthermore, we show that the generation of peptide–MHC class II complexes is controlled by TLRs in a strictly phagosome-autonomous manner.

During infection, dendritic cells (DCs) encounter and phagocytose both microbial and apoptotic cells. Without additional regulatory mechanisms, this could result in immunogenic presentation of self antigens. Here we investigated whether DCs are capable of distinguishing between the two sources of antigens for their selective presentation by MHC II.

Bone-marrow-derived DCs phagocytose apoptotic cells and deliver them into LAMP1- (lysosomal-associated membrane protein 1), MHCII-positive compartments (Fig. 1a and Supplementary Figs 1–3a). Because apoptotic cells lack TLR ligands and do not induce DC maturation (which involves surface expression of MHC II and co-stimulatory molecules, and production of cytokines)^{12,13}, MHC II remains intracellular, colocalizing with LAMP1 (Fig. 1a and Supplementary Figs 3a, 4). In contrast, phagocytosis of microbial cells, simultaneous phagocytosis of apoptotic and microbial cells, and phagocytosis of apoptotic cells accompanied by treatment with the bacterial endotoxin lipopolysaccharide (LPS) all result in DC maturation (Fig. 1b–d and Supplementary Figs 3b–d, 4 and 5). These differences reflect the transport of MHCII from late endosomal, LAMP1-positive compartments to the plasma membrane, a process that occurs upon stimulation of DCs with TLR ligands^{14,15}. We also tested the effect of phagocytosis of apoptotic cells modified to contain the TLR4 ligand LPS. Live cells efficiently internalize and retain LPS^{16,17}, and phagocytosis of apoptotic cells generated from LPS-treated B cells (LPS B-cell blasts) results in TLR4-dependent DC maturation (Fig. 1e and Supplementary Figs 3e, 4 and 5).

DCs segregate microbial and apoptotic cells into distinct phagosomes, as determined by colocalization and LAMP1 staining to

demarcate individual phagosome membranes (Fig. 1f). Given that simultaneous phagocytosis of bacterial and apoptotic cells results in plasma membrane expression of MHC II, we asked whether peptides presented by these MHC II complexes were derived from bacterial or apoptotic cells, or both. To address this question, we used two model antigens, ovalbumin (OVA) and E α protein (EAP) expressed in either bacterial or apoptotic cells, and monitored their presentation by DCs in context of the MHC II complex I-A^b using T cells derived from T-cell-receptor (TCR)-transgenic mice with TCR specificity for either OVA (OT-II) or EAP (1H3.1)^{18,19}. To assess simultaneous phagocytosis, apoptotic cells were labelled with carboxyfluorescein-diacetate succinimidyl ester (CFSE) and phagocytosed along with EAP-positive *Escherichia coli*. Almost 60% of DCs positive for CFSE stained with the YAc antibody, which detects EAP_{52–68}–I-A^b complexes²⁰ generated only after DCs have phagocytosed EAP-positive *E. coli* (Supplementary Fig. 6).

Sorted CFSE-positive DCs were used for stimulation of naive CD4⁺ T cells isolated from OT-II and 1H3.1 transgenic mice. Based on the current understanding of the mechanisms of MHC II antigen presentation^{21,22}, mature DCs should present both exogenous antigens equally well. However, we observed that bone-marrow-derived DCs containing both microbial and apoptotic cells presented the antigen derived from bacterial cells (EAP) but not the antigen from apoptotic cells (OVA) (Fig. 2a). This was also the case when EAP was from the apoptotic cells and OVA was from the bacteria (Fig. 2a), suggesting that the differences in antigen presentation were not due to differences in peptide–MHC II binding affinities, or the affinities of MHC II/peptide–TCR interactions. Similar results were obtained with splenic DCs (Fig. 2b and Supplementary Fig. 7). Antigens derived from apoptotic cells did not generate TCR ligands despite TLR engagement through simultaneous phagocytosis of bacteria, whereas antigens derived from bacteria were presented regardless of concomitant phagocytosis of apoptotic cells.

When apoptotic LPS B-cell blasts were phagocytosed by DCs, EAP was efficiently presented to 1H3.1 T cells (Fig. 2a, b and Supplementary Fig. 7), demonstrating that EAP was available for presentation when derived from apoptotic cells that contained a TLR ligand. Surface-staining of DCs with YAc consistently showed formation of this epitope only after phagocytosis of EAP-positive apoptotic LPS blasts (Supplementary Fig. 8). Antigen presentation and DC maturation after phagocytosis of apoptotic LPS B-cell blasts were completely dependent on stimulation of TLR4 in DCs (Fig. 2c and Supplementary Figs 4, 5), as apoptotic cells generated from anti-immunoglobulin-stimulated B cells induced neither antigen presentation nor DC maturation (Fig. 2c and Supplementary Fig. 4). In addition, MHCII presentation of antigen derived from apoptotic LPS B-cell blasts was not due to increased levels of EAP on these cells compared to apoptotic B cells that were not treated with LPS (Supplementary Fig. 9).

¹Section of Immunobiology, Howard Hughes Medical Institute, Yale University School of Medicine, 300 Cedar Street, New Haven, Connecticut 06520, USA. [†]Present address: Center for Immunobiology, Mount Sinai School of Medicine, 1425 Madison Avenue, New York, New York 10029, USA.

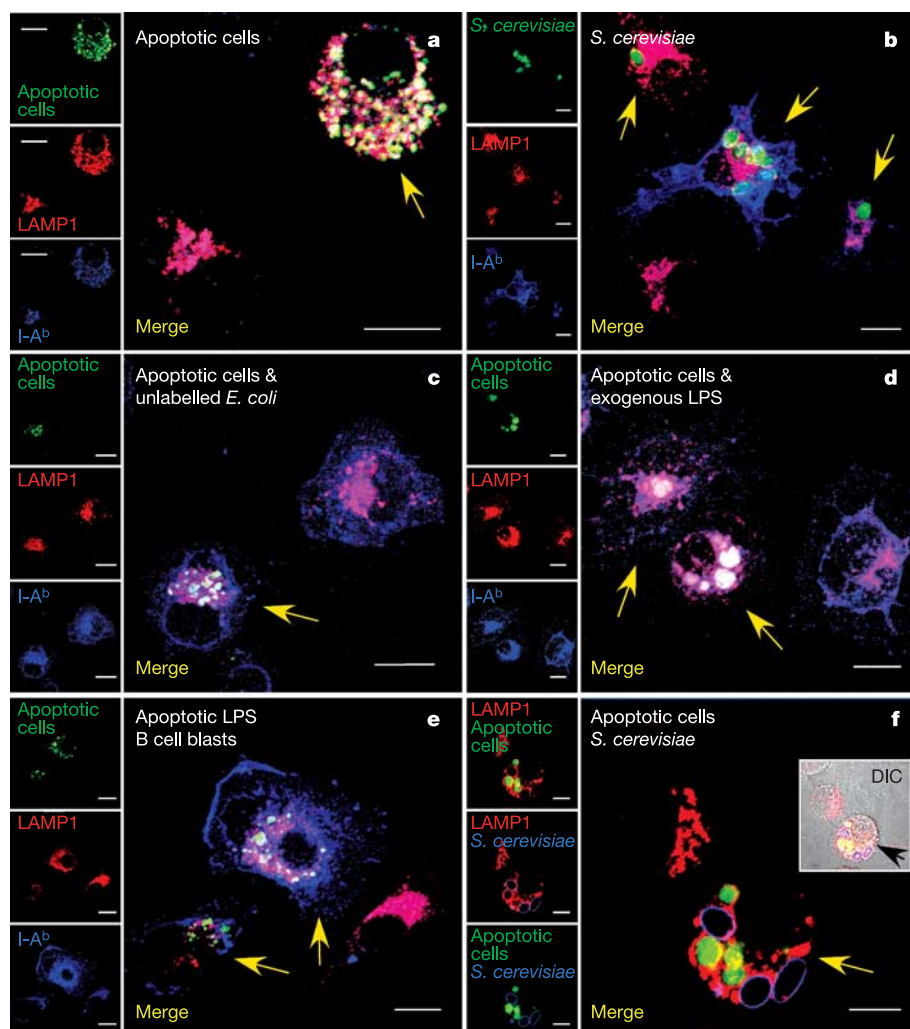


Figure 1 | Apoptotic cells phagocytosed into MHC-II⁺, LAMP-1⁺ compartments cause DC maturation only when accompanied by a TLR signal. CD11c⁺ cells were stained for intracellular LAMP1 and I-A^b (KH74 antibody) 6 h after phagocytosis to assess colocalization. Magenta indicates lysosomal MHC II, blue indicates plasma membrane MHC II (I-A^b), and white indicates cargo within MHC II⁺, LAMP1⁺ compartments. *S. cerevisiae* were stained with opsonizing reagent. Apoptotic cells were labelled with CFSE. **a–f**, Confocal micrographs showing phagocytosis of apoptotic A20 cells (**a**), *S. cerevisiae* (**b**), apoptotic A20 cells and unlabelled *E. coli* (**c**), apoptotic A20 cells and exogenous LPS (**d**), apoptotic A20 LPS blasts (**e**) or apoptotic A20 cells and *S. cerevisiae* (**f**). Arrows indicate DCs containing cargo. DIC, differential interference contrast. Scale bars, 10 μm.

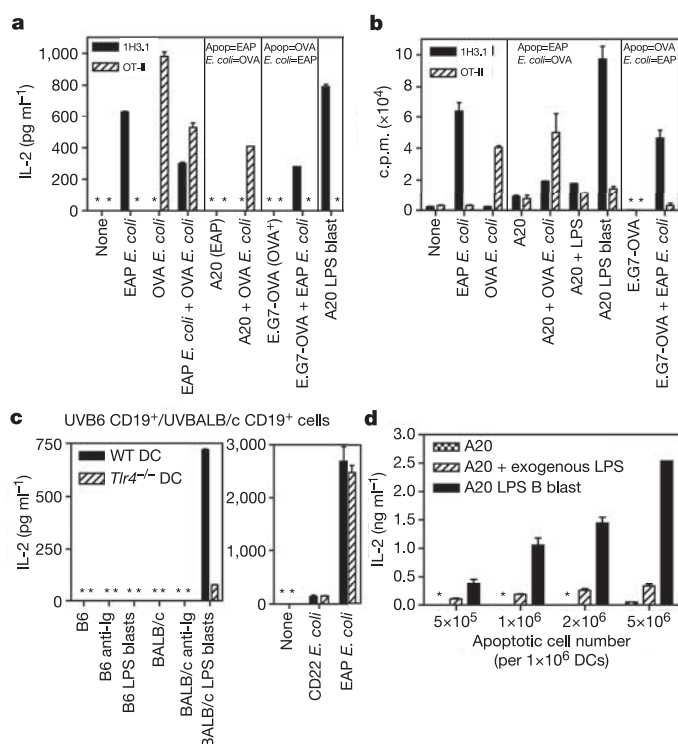


Figure 2 | Antigens from apoptotic cells are not presented by MHC II unless they carry a TLR ligand. **a**, Levels of the T-cell cytokine interleukin (IL)-2, secreted upon T-cell activation in response to bone-marrow-derived CD11c⁺ DCs, and measured at 24 h in culture supernatants. **b**, Proliferation of OT-II and 1H3.1 CD4 T cells at 72 h in response to splenic CD11c⁺ DCs. Phagocytosed cargos are indicated on the x-axes: *E. coli* expressed either OVA or EAP; E.G7-OVA are EL4 cells expressing OVA. **c**, **d**, IL-2 secretion by 1H3.1 CD4 T cells at 24 h in response to wild-type (WT) or *Tlr4*^{-/-} DCs (**c**), or wild-type DCs (**d**) as indicated. UVB6 CD19⁺ and UVBALB/c CD19⁺ cells are ultraviolet-light-irradiated apoptotic B cells derived from C57BL/6 or BALB/c mice, respectively. EAP is expressed in BALB/c but not C57BL/6 mice³⁰. Asterisks indicate 'not detected'. Error bars in **a–d** represent s.d.

As a further control for antigen amount and availability, DCs were given increasing numbers of apoptotic cells (up to 3×10^7 apoptotic cells per 1×10^6 DCs, data not shown), without an improvement in antigen presentation (Fig. 2d). Finally, phagocytosis of apoptotic cells with simultaneous stimulation of DCs by exogenous LPS failed to restore MHC II presentation (Fig. 2d) despite LPS-induced DC maturation, demonstrating an uncoupling between DC maturation and antigen presentation (Supplementary Figs 4, 5). As a further test of the contribution of LPS-induced co-stimulation, we used co-stimulation-independent 1H3.1 T-cell hybridomas²⁰ with similar results (Supplementary Fig. 10). Collectively, these results imply that presentation of particulate exogenous antigens by MHC II is determined by their origin, which in turn is determined by the presence or absence of TLR ligands within phagocytosed cargo.

We next used antigen-conjugated microspheres that allowed us to directly control the size of phagocytosed cargo, the amount and form of antigen phagocytosed by DCs, and the presence or absence of a TLR ligand on the cargo. We conjugated LPS-free hen egg lysozyme (HEL) to the surface of inert microspheres and adsorbed half of these microspheres with LPS, leaving the other half untreated. Only HEL derived from HEL/LPS microspheres that engaged TLR signalling (Supplementary Fig. 11) was efficiently presented to and activated I-A^k-restricted HEL-specific CD4⁺ T cells derived from 3A9 TCR transgenic mice (Fig. 3a). The simultaneous addition of exogenous LPS again did not result in appreciable MHC II presentation (Fig. 3a). Furthermore, the presence of LPS on different microspheres phagocytosed simultaneously by DCs along with HEL microspheres did not increase MHC II presentation (Fig. 3b). Surface staining with the HEL_{48–62}-I-A^k-specific monoclonal antibody C4H3²³ showed expression of specific HEL peptide–MHC II complexes only on

DCs that had phagocytosed HEL/LPS microspheres (Fig. 3c, d and Supplementary Fig. 12). Consistent with immunofluorescence observations, flow cytometry showed that only HEL/LPS microspheres induced cell-surface expression of MHC II in most cells (Fig. 3e).

These results demonstrate the existence of TLR-dependent and phagosome-autonomous selection of antigens for presentation by DCs. To investigate the mechanism of this process, we used enriched phagosomes from DCs that phagocytosed two types of microspheres (magnetic or plain), where one was TLR ligand-positive and the other TLR ligand-negative. DCs were given plain HEL/LPS microspheres and magnetic HEL microspheres, or vice versa. The population of magnetic phagosomes was separated from DC homogenates, and the population of plain phagosomes further enriched on discontinuous sucrose density gradients²⁴ and then biochemically characterized (Supplementary Fig. 13). Lysosomal β -hexosaminidase²⁵ and LAMP2 were confined to gradient fractions 10–11 (F10 fractions). H-2M α and LAMP2 were enriched in 2- μ g F10 fractions compared to 20- μ g post-nuclear supernatants, and were comparable in phagosomes that did or did not contain TLR ligands (Fig. 4a). In contrast, no enrichment of the Golgi matrix protein GM-130, or endoplasmic reticulum components like calnexin, was detected in F10 fractions (Fig. 4a). Notably, the p31 isoform of MHC II-associated invariant chain (Ii) was consistently degraded in HEL/LPS phagosomes, but persisted in HEL phagosomes despite the concomitant presence of LPS in magnetic phagosomes within those cells (Fig. 4a). This pattern of Ii degradation was also observed in phagosomes enriched from DCs that had phagocytosed HEL or HEL/LPS microspheres alone (Supplementary Fig. 13e). These data suggest that Ii p31 degradation is confined predominantly to phagosomes that contain TLR ligands.

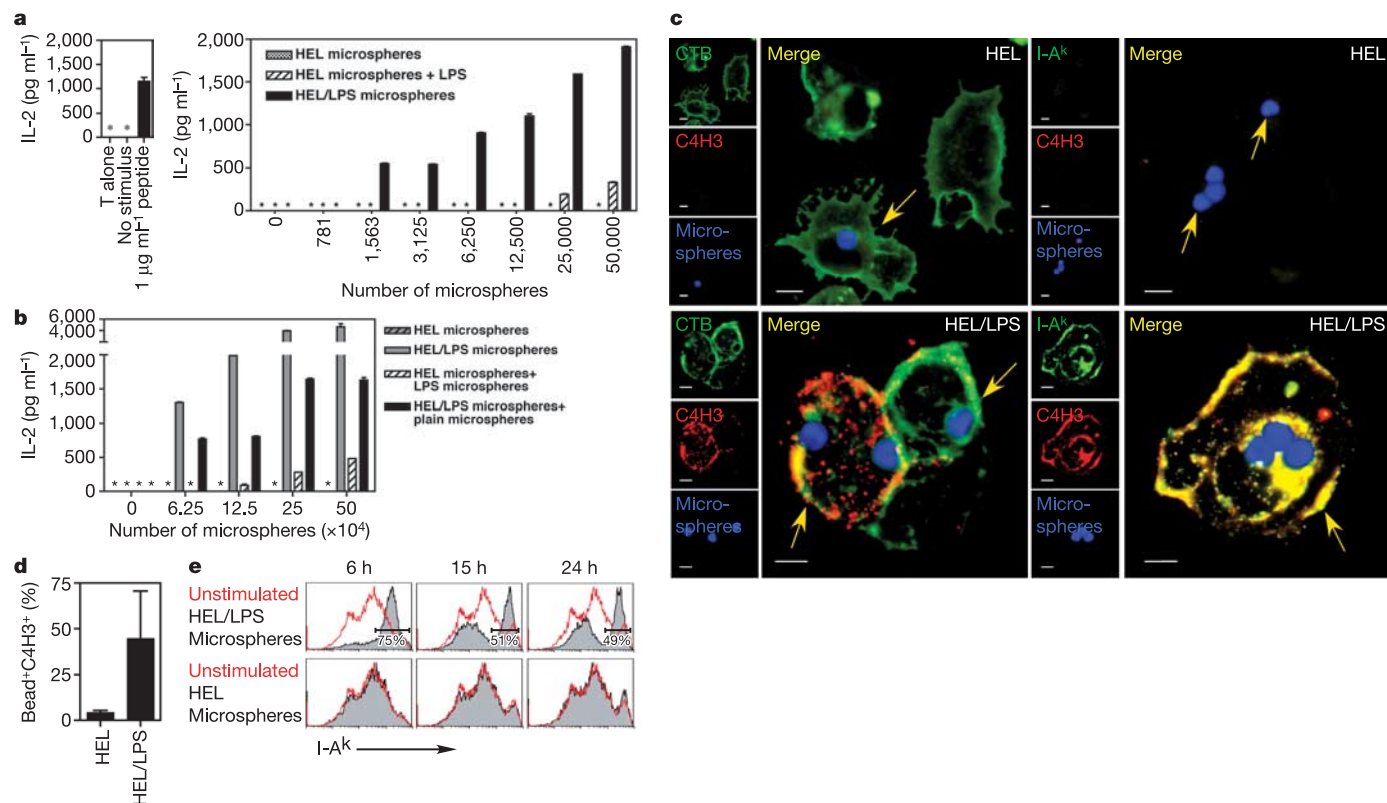


Figure 3 | MHC II presentation of phagocytosed particulate antigens is determined by the ability of cargo to engage TLRs. **a**, **b**, IL-2 secretion by 3A9 CD4 cells stimulated by DCs given LPS and HEL_{48–62} (**a**, left), microspheres as indicated (right), or in combination with LPS-adsorbed or plain microspheres (**b**). Asterisks indicate 'not detected'. **c**, Deconvoluted micrographs representing extended focus images of several Z-images show

surface CTB or I-A^k(11-5.2) with surface C4H3 on CD11c⁺ DCs. Microspheres are blue, arrows show DCs with microspheres. Scale bars, 5 μ m. **d**, **e**, Quantification (**d**) and flow cytometric analysis (**e**) of cell-surface I-A^k(11-5.2) on CD11c⁺ DCs after phagocytosis of HEL or HEL/LPS microspheres. Error bars in **a**, **b**, **d** represent s.d.

Accordingly, the levels of mature SDS-stable $\alpha\beta$ MHC II dimers were highly enriched in HEL/LPS phagosomes, in contrast to HEL phagosomes (Fig. 4a). The mature form of cathepsin L was enriched in phagosomes compared to post-nuclear supernatants before enrichment, but the levels of cathepsin L, S, B or D were similar between HEL and HEL/LPS phagosomes (Fig. 4a). It is likely that the activities, rather than the recruitment, of these enzymes are regulated by TLR signalling.

Does regulation of antigen processing contribute to differential antigen presentation by DCs? We bypassed the requirement for lysosomal proteolysis by testing MHC II presentation of a pre-processed immunogenic peptide (OVA amino acids 323–339) conjugated to microspheres with or without LPS. Presentation of OVA peptide to OT-II T cells was only detected when LPS was also adsorbed to microspheres (Fig. 4b). This shows that control of antigen processing does not account for TLR-dependent antigen selection. Rather, progressive proteolysis of Ii to CLIP²⁶ seems to be one of the TLR-regulated, phagosome-autonomous steps that contributes to antigen selection for presentation. H2-M-mediated CLIP exchange for antigenic peptide is unlikely to have a principal role, as this step occurs after Ii processing²⁷, and we did not observe significant differences in H-2M recruitment to phagosomes with or without LPS. However, additional regulatory steps are likely to exist, and indeed, TLR signalling controls multiple steps of antigen handling and presentation, including antigen uptake^{28,29}.

We next prepared microspheres conjugated to either LPS-free OVA (OVA peptide) or HEL, and either adsorbed with LPS or left untreated. Microspheres with these two antigens were simultaneously given to DCs in a combination in which only one antigen was associated with LPS. We found that the antigen conjugated to LPS microspheres was strongly favoured and predominated MHC II presentation (Fig. 4c). HEL did not contain residual contaminating OVA, as no OT-II response was observed to HEL or HEL/LPS microspheres alone (Fig. 4d). It was important to remove the bulk of soluble unbound protein and LPS, to limit presentation to particulate rather than soluble antigen (Supplementary Fig. 14a, b), and to minimize minor contributions that exogenous unbound LPS may have on presentation (Figs 2d, 3a). However, residual soluble HEL and LPS remained, as assessed by testing microsphere supernatants for presentation and cytokine release by DCs (Supplementary Fig. 14c and data not shown, respectively).

To directly observe phagosome-autonomous formation of immunogenic peptide–MHC II complexes at the single-cell level, we monitored the intracellular formation of the C4H3 epitope on phagosomes. Using two different sizes of HEL-conjugated microspheres differentially adsorbed with LPS, we were able to discriminate HEL from HEL/LPS phagosomes by confocal microscopy, and found preferential formation of HEL_{48–62}–I-A^k complexes on HEL/LPS phagosomes regardless of the size of cargo (Fig. 4e and Supplementary Videos S1 and S2). Staining for cholera toxin B

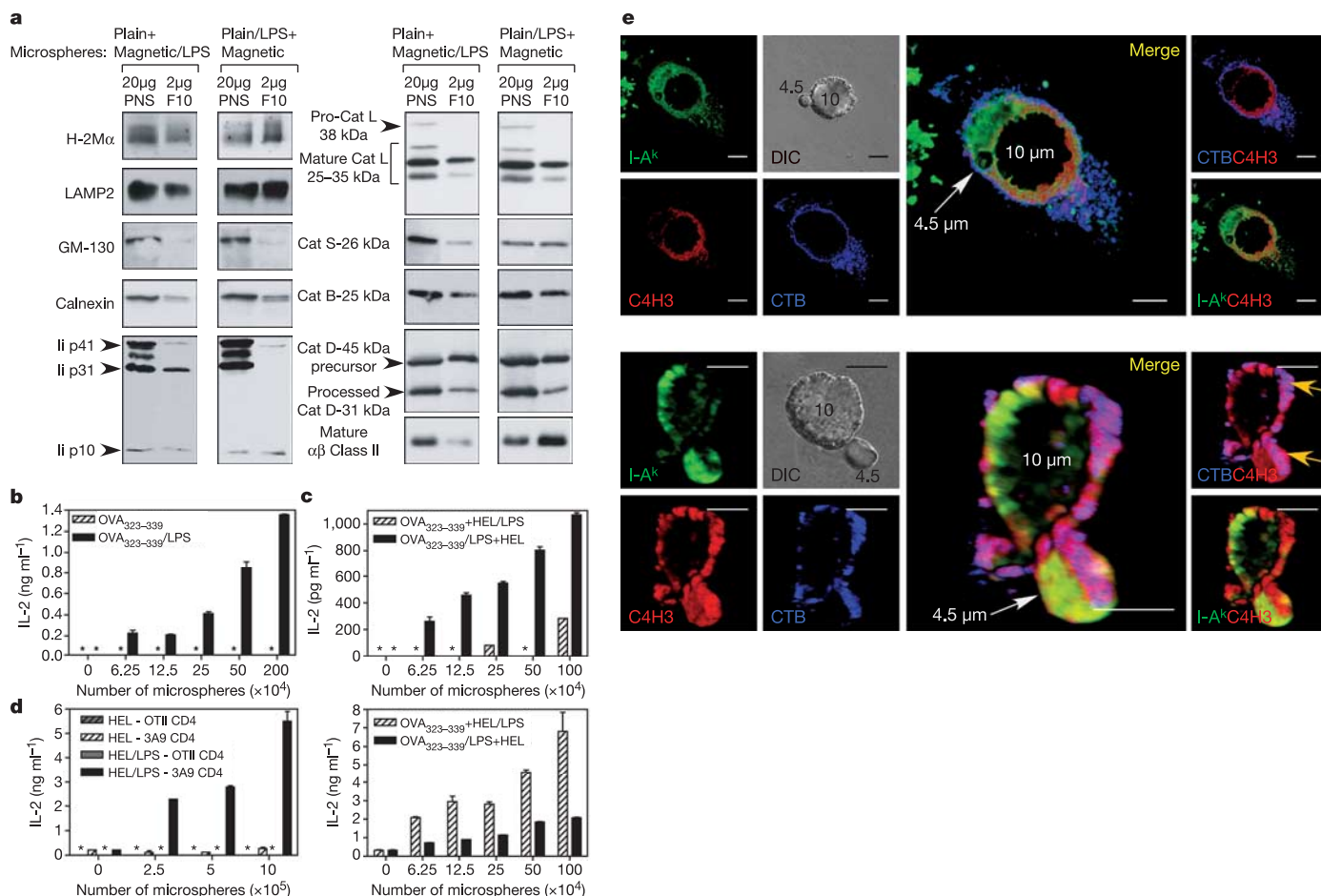


Figure 4 | TLR engagement by phagocytosed cargo controls phagosome-autonomous MHC II presentation. **a**, Western blots of DC post-nuclear supernatant (PNS) and F10 phagosomes after exclusion of magnetic microspheres phagocytosed simultaneously with plain microspheres. Cat, cathepsin. **b**, IL-2 secretion by OT-II cells in response to DCs given OVA_{323–339} microspheres. **c**, IL-2 secretion by OT-II (top) or 3A9 cells (bottom) stimulated with F₁ (CBA/J × C57BL/6J) DCs. **d**, IL-2 secretion by

either OT-II or 3A9 cells in response to F₁ DCs given either HEL or HEL/LPS microspheres. Asterisks indicate 'not detected'. **e**, Three-dimensional reconstructions of DCs stained for intracellular C4H3, I-A^k_(11–5.2) and surface CTB 5 h after simultaneous phagocytosis of 10 μ m HEL/LPS and 4.5 μ m HEL microspheres (top) or 10 μ m HEL and 4.5 μ m HEL/LPS microspheres (bottom). Scale bars, 5 μ m. Error bars in **b–d** represent s.d.

subunit (CTB) before permeabilization allowed differentiation between intracellular and surface C4H3. In lower panels of Fig. 4e, C4H3 colocalized with I-A^k inside 4.5- μ m HEL/LPS phagosomes (yellow upon merge), whereas C4H3 that followed the contour around the 10- μ m microsphere colocalized with the plasma membrane marker CTB and was not phagosomal (Fig. 4e, magenta upon merge, indicated by arrows).

Collectively, these experiments demonstrate that TLRs control the generation of TCR ligands derived from phagocytosed cargo in a phagosome-autonomous manner. Using this mechanism, DCs can classify different sources of antigens as self or non-self at the subcellular level, resulting in compartmentalized generation of peptide-MHC II complexes where the contents of phagosomes derived from microbial pathogens are preferentially presented in the context of co-stimulation. This TLR-based mechanism of discriminating phagosomal contents would ensure that upon simultaneous phagocytosis of self and non-self, antigens derived from self are excluded from the pool of MHC II transported to the plasma membrane upon concomitant TLR engagement by microbial pathogens. An important functional consequence of compartmentalized TLR-mediated control of phagosome maturation²⁸ in DCs is the selection of antigenic cargo for MHC II presentation, which, together with TLR-induced expression of co-stimulatory signals, contributes to peripheral self/non-self discrimination. It will be important to determine how presentation of antigens derived from apoptotic cells by tolerogenic DCs occurs *in vivo*, and whether antigen presentation is regulated differently in various subsets of dendritic cells.

METHODS

Detailed descriptions of reagents and experimental methods can be found in the Supplementary Methods. C4H3 antibody was provided by R. Germain, YAc and Y3JP antibodies were from C. Janeway Jr, anti-LAMP1 and LAMP2 antibodies were from I. Mellman, anti-calnexin antibody was from P. Cresswell, and anti-H-2M α antibody was from J. Monaco. E.G7-OVA cells and pAcNeo OVA were provided by M. Bevan. 1H3.1 TCR transgenic mice (and the 1H3.1 T-cell hybridoma) were a gift from C. Janeway Jr, 3A9 TCR transgenic mice were from I. Mellman, and *Tlr4*^{-/-} mice were from S. Akira. C57BL/6J, BALB/c, CBA/J and F₁ (CBA/J $\times$ C57BL/6J) mice were purchased from the Jackson Laboratory.

Received 8 November 2005; accepted 23 January 2006.

Published online 19 February 2006.

- Greenberg, S. & Grinstein, S. Phagocytosis and innate immunity. *Curr. Opin. Immunol.* **14**, 136–145 (2002).
- Ravichandran, K. S. "Recruitment signals" from apoptotic cells: invitation to a quiet meal. *Cell* **113**, 817–820 (2003).
- Savill, J., Dransfield, I., Gregory, C. & Haslett, C. A blast from the past: clearance of apoptotic cells regulates immune responses. *Nature Rev. Immunol.* **2**, 965–975 (2002).
- Henson, P. M., Bratton, D. L. & Fadok, V. A. Apoptotic cell removal. *Curr. Biol.* **11**, R795–R805 (2001).
- Iwasaki, A. & Medzhitov, R. Toll-like receptor control of the adaptive immune responses. *Nature Immunol.* **5**, 987–995 (2004).
- Steinman, R. M., Hawiger, D. & Nussenzweig, M. C. Tolerogenic dendritic cells. *Annu. Rev. Immunol.* **21**, 685–711 (2003).
- Larsson, M., Fonteneau, J. F. & Bhardwaj, N. Dendritic cells resurrect antigens from dead cells. *Trends Immunol.* **22**, 141–148 (2001).
- Steinman, R. M., Turley, S., Mellman, I. & Inaba, K. The induction of tolerance by dendritic cells that have captured apoptotic cells. *J. Exp. Med.* **191**, 411–416 (2000).
- Henson, P. M., Bratton, D. L. & Fadok, V. A. The phosphatidylserine receptor: a crucial molecular switch? *Nature Rev. Mol. Cell Biol.* **2**, 627–633 (2001).
- Trombetta, E. S. & Mellman, I. Cell biology of antigen processing *in vitro* and *in vivo*. *Annu. Rev. Immunol.* **23**, 975–1028 (2005).
- Banchereau, J. & Steinman, R. M. Dendritic cells and the control of immunity. *Nature* **392**, 245–252 (1998).
- Sauter, B. *et al.* Consequences of cell death: exposure to necrotic tumor cells, but not primary tissue cells or apoptotic cells, induces the maturation of immunostimulatory dendritic cells. *J. Exp. Med.* **191**, 423–434 (2000).
- Gallucci, S., Lolkema, M. & Matzinger, P. Natural adjuvants: endogenous activators of dendritic cells. *Nature Med.* **5**, 1249–1255 (1999).
- Turley, S. J. *et al.* Transport of peptide-MHC class II complexes in developing dendritic cells. *Science* **288**, 522–527 (2000).
- Chow, A., Toomre, D., Garrett, W. & Mellman, I. Dendritic cell maturation triggers retrograde MHC class II transport from lysosomes to the plasma membrane. *Nature* **418**, 988–994 (2002).
- Thieblemont, N. & Wright, S. D. Transport of bacterial lipopolysaccharide to the Golgi apparatus. *J. Exp. Med.* **190**, 523–534 (1999).
- Latz, E. *et al.* Lipopolysaccharide rapidly traffics to and from the Golgi apparatus with the Toll-like receptor 4-MD-2-CD14 complex in a process that is distinct from the initiation of signal transduction. *J. Biol. Chem.* **277**, 47834–47843 (2002).
- Barnden, M. J., Allison, J., Heath, W. R. & Carbone, F. R. Defective TCR expression in transgenic mice constructed using cDNA-based α - and β -chain genes under the control of heterologous regulatory elements. *Immunol. Cell Biol.* **76**, 34–40 (1998).
- Viret, C. & Janeway, C. A., Jr. Functional and phenotypic evidence for presentation of E α 52–68 structurally related self-peptide(s) in I-E α -deficient mice. *J. Immunol.* **164**, 4627–4634 (2000).
- Rudensky, A., Rath, S., Preston-Hurlburt, P., Murphy, D. B. & Janeway, C. A., Jr. On the complexity of self. *Nature* **353**, 660–662 (1991).
- Watts, C. The exogenous pathway for antigen presentation on major histocompatibility complex class II and CD1 molecules. *Nature Immunol.* **5**, 685–692 (2004).
- Boes, M., Cuvillier, A. & Ploegh, H. Membrane specializations and endosome maturation in dendritic cells and B cells. *Trends Cell Biol.* **14**, 175–183 (2004).
- Zhong, G., Reis e Sousa, C. & Germain, R. N. Production, specificity, and functionality of monoclonal antibodies to specific peptide-major histocompatibility complex class II complexes formed by processing of exogenous protein. *Proc. Natl Acad. Sci. USA* **94**, 13856–13861 (1997).
- Luhrmann, A. & Haas, A. A method to purify bacteria-containing phagosomes from infected macrophages. *Methods Cell Sci.* **22**, 329–341 (2000).
- Storrie, B. & Madden, E. A. Isolation of subcellular organelles. *Methods Enzymol.* **182**, 203–225 (1990).
- Cresswell, P. Invariant chain structure and MHC class II function. *Cell* **84**, 505–507 (1996).
- Denzin, L. K. & Cresswell, P. HLA-DM induces CLIP dissociation from MHC class II $\alpha\beta$ dimers and facilitates peptide loading. *Cell* **82**, 155–165 (1995).
- Blander, J. M. & Medzhitov, R. Regulation of phagosome maturation by signals from Toll-like receptors. *Science* **304**, 1014–1018 (2004).
- West, M. A. *et al.* Enhanced dendritic cell antigen capture via Toll-like receptor-induced actin remodeling. *Science* **305**, 1153–1157 (2004).
- Mathis, D. J., Benoist, C., Williams, V. E. II, Kanter, M. & McDevitt, H. O. Several mechanisms can account for defective E α gene expression in different mouse haplotypes. *Proc. Natl Acad. Sci. USA* **80**, 273–277 (1983).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to I. Mellman, M. Bevan, J. Monaco and R. Germain for reagents, and E. Kopp for the flagellin-OVA fusion construct. We thank P. Cresswell and A. Rudensky for reagents, suggestions and critical reading of the manuscript. We are indebted to I. Mellman, E.S. Trombetta, L. Delamarre and J. Kagan for discussions; M. Pypaert and C. Marks for electron microscopy; C. Chalouni and R. Montgomery for confocal microscopy; A. Hartley for 3D reconstructions; D. Amsen, A. Giodini, R. Teal, T. Horng, T. Potseleuva, C. Annicelli, S. Holley, R. Couture, D. Matosian, T. Taylor, G. Tokmoulina, J. Shanabrough, M. Blander and S.J. Blander for assistance and support. J.M.B. is supported by an Arthritis Investigator award. R.M. is supported by the Howard Hughes Medical Institute and the NIH.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.M. (ruslan.medzhitov@yale.edu).

Omega-3 and omega-6 fatty acids stimulate cell membrane expansion by acting on syntaxin 3

Frédéric Darios¹ & Bazbek Davletov¹

Growth of neurite processes from the cell body is the critical step in neuronal development and involves a large increase in cell membrane surface area¹. Arachidonic-acid-releasing phospholipases are highly enriched in nerve growth cones and have previously been implicated in neurite outgrowth^{2,3}. Cell membrane expansion is achieved through the fusion of transport organelles with the plasma membrane⁴; however, the identity of the molecular target of arachidonic acid has remained elusive. Here we show that syntaxin 3 (STX3), a plasma membrane protein, has an important role in the growth of neurites, and also serves as a direct target for omega-6 arachidonic acid. By using syntaxin 3 in a screening assay, we determined that the dietary omega-3 linolenic and docosahexaenoic acids can efficiently substitute for arachidonic acid in activating syntaxin 3. Our findings provide a molecular basis for the previously established action of omega-3 and omega-6 polyunsaturated fatty acids in membrane expansion at the growth cones, and represent the first identification of a single effector molecule for these essential nutrients.

Pheochromocytoma neuroendocrine cells (PC12 cells) readily emit neurites when treated with nerve growth factor (NGF), presenting a convenient model of cell membrane expansion^{5,6}. Neurite outgrowth involves the activation of phospholipases that release soluble fatty acids such as the ubiquitous arachidonic acid^{2,3,5,7}. Indeed, co-treatment of PC12 cells with NGF and phospholipase A2 (PLA2) had a strong stimulatory effect on the length of neurites (Fig. 1a). A 30-min pretreatment of PC12 cells with arachidonic acid, before the addition of NGF, was also able to enhance neurite outgrowth without affecting cell viability (Fig. 1a and Supplementary Fig. S1). The fusion of plasmalemmal precursor vesicles into the cell surface membrane requires SNARE (soluble N-ethylmaleimide-sensitive-factor attachment protein receptor) fusion proteins^{8,9} and, by definition, involves lipids^{10,11}. We recently reported that Munc18-regulated syntaxin 1 (STX1), a plasma membrane SNARE specific for neuroendocrine cells, is sensitive to arachidonic acid¹². Therefore, we tested whether knockdown of syntaxin 1 would interfere with the ability of PC12 cells to grow neurites. Transfection of PC12 cells with an *Stx1*-specific short interfering (si)RNA resulted in the down-regulation of syntaxin 1 in approximately 20% of cells, as judged by immunostaining (data not shown). These cells, however, were still capable of neurite outgrowth when stimulated by NGF (Fig. 1b and Supplementary Fig. S2). This result is consistent with observations that syntaxin 1 is specifically involved in fast calcium-triggered exocytosis of neurotransmitters and not in axonal growth¹³. Hence, we considered whether another plasma membrane syntaxin is (1) responsible for neurite outgrowth and (2) responsive to arachidonic acid in a similar manner to syntaxin 1.

Among the plasma membrane syntaxins 1–4, syntaxin 3 is conspicuously enriched in rapidly growing cells¹⁴, is present in neuronal growth cones (Supplementary Fig. S3), and is a target for botulinum toxin C, which potently blocks axonal growth^{15,16}. Therefore, we

tested the effect of siRNA-mediated knockdown of syntaxin 3 on neurite outgrowth. Figure 1c shows that PC12 cells lacking syntaxin 3 were unable to grow neurites. The absence of syntaxin 3 had as profound an effect on embryonic neurons as on PC12 cells (Fig. 1d),

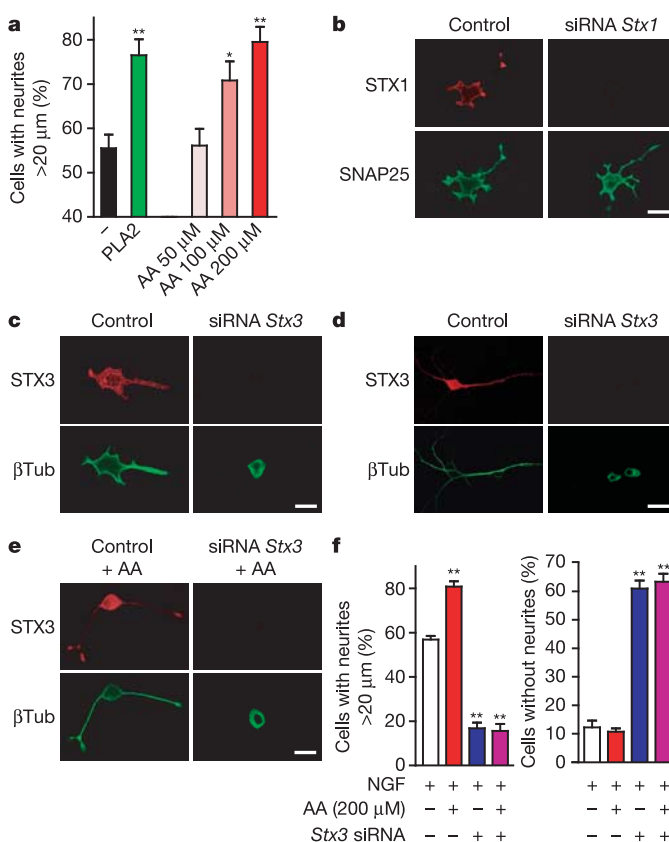


Figure 1 | Involvement of syntaxin 3 in neurite outgrowth. **a**, PLA2 ($10 \mu\text{g ml}^{-1}$) and arachidonic acid (AA) upregulate NGF-induced neurite outgrowth in PC12 cells. The percentage of PC12 cells carrying neurites longer than $20 \mu\text{m}$ is shown. Increasing concentrations of arachidonic acid (50 – $200 \mu\text{M}$) were tested. **b**, Syntaxin 1 (STX1) depletion in siRNA-treated PC12 cells does not interfere with neurite outgrowth. Example images of control and siRNA-treated cells immunostained for STX1 and SNAP25 are shown. **c**, Syntaxin 3 (STX3) depletion in siRNA-treated PC12 cells blocks neurite outgrowth. Immunostaining of syntaxin 3 and β III-tubulin (β Tub) is shown. **d**, Syntaxin 3 depletion also abolishes branching in cultured hippocampal neurons. **e**, Syntaxin 3 depletion in siRNA-treated PC12 cells abolishes neurite outgrowth induced by NGF and $200 \mu\text{M}$ arachidonic acid. **f**, Effect of syntaxin 3 depletion on the percentage of PC12 cells either carrying neurites longer than $20 \mu\text{m}$ (left panel) or lacking neurites altogether (right panel). Error bars are s.e.m.; *, $P < 0.05$; **, $P < 0.01$. Scale bars, $20 \mu\text{m}$.

¹MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

confirming the physiological relevance of this syntaxin isoform. Notably, even in cells that were co-treated with arachidonic acid and NGF, the downregulation of syntaxin 3 had a powerful inhibitory effect on neurite outgrowth (Fig. 1e, f).

In order to probe the sensitivity of syntaxin 3 to arachidonic acid, we first determined whether it can interact with SNAP25 (synaptosomal-associated protein of 25 kDa), a syntaxin partner implicated in neurite outgrowth¹⁷. Co-immunostaining of the two proteins demonstrated strong co-localization of SNAP25 with syntaxin 3 in the growth cones of neurites (Fig. 2a). Immunoprecipitation of syntaxin 3 from resting PC12 cells led to a pull-down of SNAP25, but not SNAP23 (Fig. 2b), highlighting the specificity of STX3–SNAP25 interaction. Remarkably,

NGF treatment enhanced the interaction between syntaxin 3 and SNAP25, as evidenced by either syntaxin 3 or SNAP25 pull-down using specific antibodies (Fig. 2b, c). Because NGF triggers a complex anabolic response⁵, which includes arachidonic acid release⁷, we tested the role of arachidonic acid itself. Quantitative immunoprecipitation demonstrated that a 30-min treatment of either resting PC12 cells or neuronal cultures with exogenous arachidonic acid was sufficient to stimulate the SNARE interaction (Fig. 2d, e). Thus, the two proteins are apparently sensitive to a rise in arachidonic acid concentration. The basal levels of STX3–SNAP25 interaction detected in the absence of NGF or arachidonic acid treatment (Fig. 2b–e) may be a reflection of growth processes already taking place in these cells.

Next, we investigated the effects of arachidonic acid in defined molecular reactions. SNAP25 was incubated for 30 min with syntaxin 3 in the absence or presence of arachidonic acid, and the reactions analysed by size-exclusion chromatography. In the absence of arachidonic acid, syntaxin 3 migrated behind SNAP25 (Fig. 3a), suggesting that (1) the two proteins do not interact under these conditions, and (2) syntaxin 3 has a more compact structure than SNAP25. However, the addition of arachidonic acid was sufficient to allow formation of the binary STX3–SNAP25 complex (Fig. 3a). A pull-down experiment using immobilized glutathione S-transferase (GST)-tagged syntaxin 3 (GST–STX3) further confirmed that SNAP25 can bind syntaxin 3 only in the presence of arachidonic acid (Fig. 3b). Titration of arachidonic acid revealed that its half-maximal effective concentration (EC_{50}) was $\sim 100 \mu\text{M}$ (Fig. 3b). By taking into account Avogadro's number, $100 \mu\text{M}$ can be defined as one molecule present in a 25-nm cube—a plausible concentration at the phospholipid bilayer where syntaxin 3 resides. As arachidonic acid is highly enriched in natural phospholipid bilayers^{18,19}, the local action of phospholipases allows its localized release in micromolar concentrations²⁰. Notably, isolated growth cones show a high turnover of arachidonic acid, with the bulk free-concentration reaching 2.5% of the total content²¹, which can be translated into a concentration range of $\sim 100 \mu\text{M}$. We tested whether PLA2 treatment of resting PC12 cells would be sufficient to enhance the STX3–SNAP25 interaction. Figure 3c shows that PLA2 treatment led to a marked increase in syntaxin 3 immunoprecipitated with SNAP25, indicating that the enzyme can release, even if at the plasma membrane, sufficient amounts of arachidonic acid to activate the SNAREs.

The above data suggest that either syntaxin 3 or SNAP25 can sense arachidonic acid. Therefore, we investigated global structural changes in the two individual proteins. Circular dichroism (CD) experiments demonstrated that syntaxin 3 responds to arachidonic acid by an increase in the alpha-helical content of the protein, whereas no changes were apparent in the case of SNAP25 (Fig. 3d). Of note was the observation that the two proteins, when mixed in the presence of arachidonic acid, showed an additional increase in alpha-helicity compared with the theoretical spectrum²², again demonstrating that the two proteins physically interact (Supplementary Fig. S4). To further dissect the action of arachidonic acid, we analysed protein intrinsic fluorescence, which monitors the micro-environment around aromatic amino-acid residues and thus provides a readout of conformational changes²³. Arachidonic acid had no effect on the SNAP25 tryptophan spectrum, but triggered drastic changes in the environment of the syntaxin 3 tyrosines, with an EC_{50} close to $100 \mu\text{M}$ (Fig. 3e). An established 1,8-anilinonaphthalene-sulphonic acid (ANS)-based assay for fatty acid binding²⁴ independently confirmed that syntaxin 3 can interact with arachidonic acid in this micromolar range (Fig. 3f).

We then analysed the ability of syntaxin 3 to form the SNARE ternary complex implicated in membrane fusion^{8,9}. We used synaptobrevin 2 (SYB2; also known as VAMP2) in this reaction because it allows formation of the SDS-resistant SNARE complex similar to several other vesicular SNAREs^{25,26}. In the absence of arachidonic acid, syntaxin 3 was unable to form the ternary SNARE complex

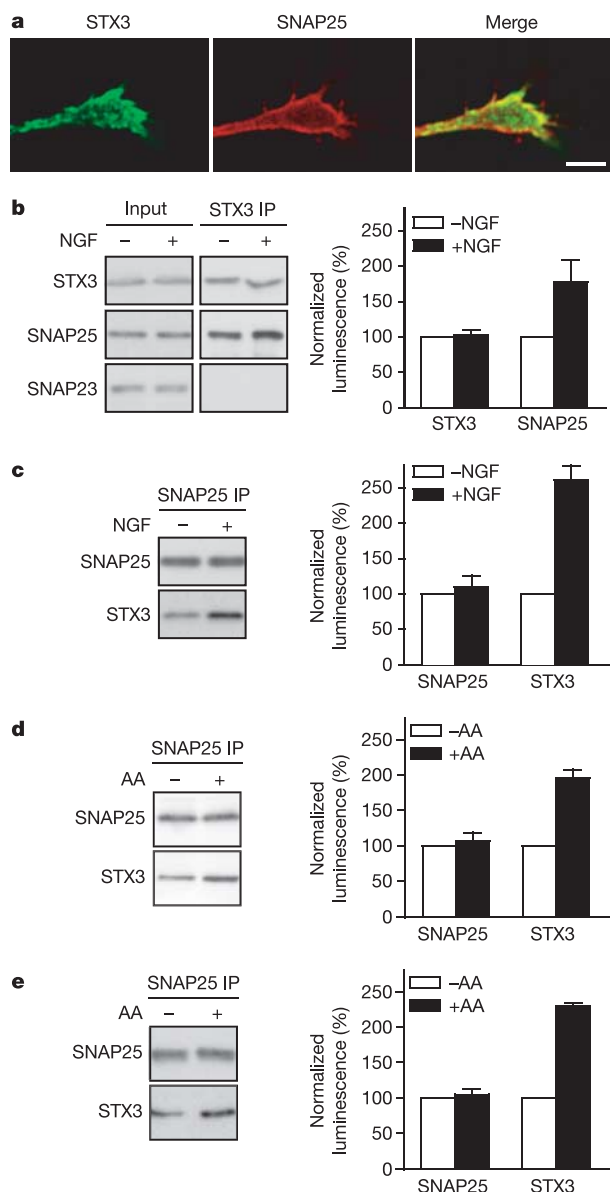


Figure 2 | Potentiation of STX3–SNAP25 interaction by arachidonic acid.

a, Double immunostaining of syntaxin 3 and SNAP25 demonstrates co-localization of the two proteins in growth cones. Scale bar, 200 nm. **b**, Syntaxin 3 immunoprecipitates (IP) with SNAP25 but not SNAP23 (left panels). NGF treatment of PC12 cells potentiates syntaxin 3 interaction with SNAP25, as quantified by chemiluminescent immunoblotting (right panel). **c**, SNAP25 pull-down confirms the enhanced STX3–SNAP25 interaction following NGF treatment. **d**, **e**, A 30-min treatment of PC12 cells (**d**) or embryonic hippocampal neurons (**e**) with $200 \mu\text{M}$ arachidonic acid, in the absence of NGF, enhances STX3–SNAP25 interaction. Error bars are s.e.m.; $n = 3$.

(Fig. 3g). Arachidonic acid alone was sufficient to overcome the inhibition of syntaxin 3 to trigger formation of the SNARE complex, with an EC_{50} of $\sim 100 \mu\text{M}$. To determine whether arachidonic acid acts only at the step of syntaxin–SNAP25 assembly, we incubated synaptobrevin with the syntaxin–SNAP25 dimer that was previously isolated by size-exclusion chromatography (Fig. 3a). Figure 3h demonstrates that when the syntaxin–SNAP25 binary complex was already assembled, subsequent synaptobrevin engagement into the

SNARE complex did not require the further presence of arachidonic acid.

Neuronal phospholipids often carry, in place of arachidonic acid, omega-3 docosahexaenoic acid^{18,19}, which is also released by PLA2 enzymes. Hence, we screened different saturated and unsaturated fatty acids (Supplementary Fig. S5), including polyunsaturated fatty acids (PUFAs), for their ability to influence syntaxin 3. Docosahexaenoic, omega-3 linolenic and omega-6 linoleic acids efficiently activated

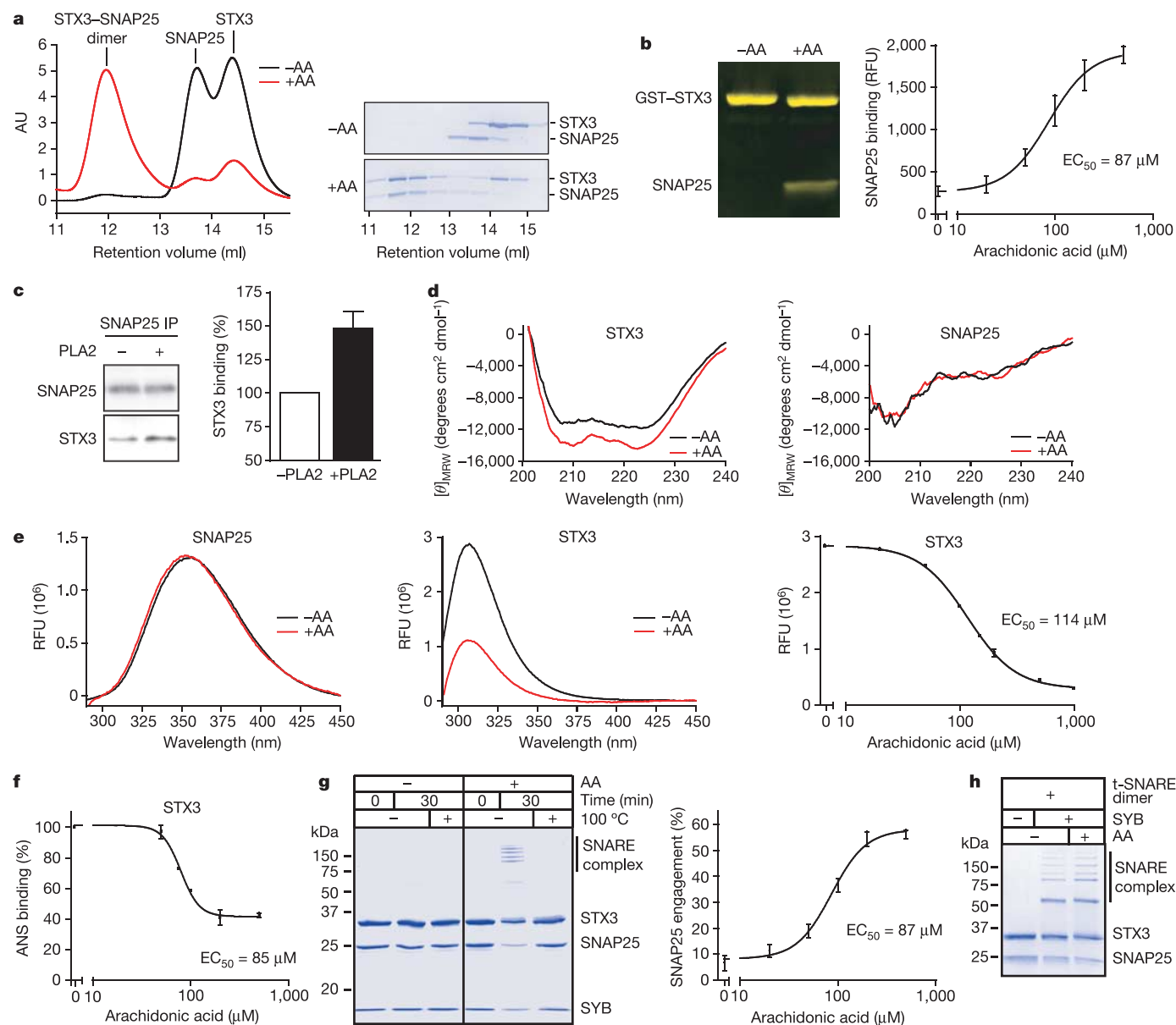


Figure 3 | Arachidonic acid activates syntaxin 3 to allow SNARE assembly. **a**, Arachidonic acid induces formation of a stable STX3–SNAP25 binary complex, as assessed by size-exclusion chromatography (left panel). The Coomassie-stained gel demonstrates co-migration of syntaxin 3 and SNAP25 in early chromatographic fractions (right panel). AU, arbitrary absorbance units at 280 nm. **b**, Sypro-orange-stained gel showing arachidonic acid requirement for SNAP25 interaction with immobilized GST–STX3 (left panel). Titration of arachidonic acid in the pull-down assay yields an $EC_{50} = 87 \mu\text{M}$ (right panel). RFU, relative fluorescence units. **c**, PLA2 treatment of resting PC12 cells potentiates syntaxin 3 interaction with SNAP25, as measured by their co-immunoprecipitation. **d**, CD spectra of syntaxin 3 (left panel) and SNAP25 (right panel) in the absence or presence of arachidonic acid. $[\theta]_{\text{MRW}}$, mean residue molar ellipticity. **e**, Intrinsic protein fluorescence spectra of SNAP25 (left panel; excitation

wavelength 280 nm) and syntaxin 3 (middle panel; excitation wavelength 280 nm) in the absence or presence of arachidonic acid. The effect of arachidonic acid titration on syntaxin 3 autofluorescence suggests an $EC_{50} = 114 \mu\text{M}$ (right panel; excitation wavelength 280 nm, emission wavelength 306 nm). **f**, ANS-based competition assay for arachidonic acid binding indicates an $EC_{50} = 85 \mu\text{M}$. **g**, Coomassie-stained gel showing that arachidonic acid is necessary for assembly of syntaxin 3, SNAP25 and synaptobrevin (SYB) into the SDS-resistant SNARE complex (left panel). Quantification of SNAP25 engagement into the SDS-resistant complex indicates an $EC_{50} = 87 \mu\text{M}$ (right panel). **h**, Coomassie-stained gel showing that arachidonic acid is not required at the step of synaptobrevin engagement with the stable STX3–SNAP25 (t-SNARE) dimer. Single-concentration experiments used $200 \mu\text{M}$ arachidonic acid. Error bars are s.e.m.; $n = 3$.

syntaxin 3 in the SNARE complex-formation assay (Fig. 4a, b). Note that linoleic and linolenic acids are plant-derived nutrient compounds essential for all mammals, and serve as precursors of arachidonic and docosahexaenoic acid, respectively^{19,27}. Next, we tested the effect of the fatty acids on the syntaxin 3 molecule alone. The intrinsic tyrosine fluorescence of syntaxin 3 showed marked changes upon addition of omega-3 linolenic and docosahexaenoic acids, but neither omega-9 mead acid, saturated fatty acids, or monounsaturated fatty acids were effective (Fig. 4c). If syntaxin 3 is indeed the essential PUFA effector molecule during neurite outgrowth, then linolenic and docosahexaenoic acids should stimulate neurite outgrowth in NGF-treated PC12 cells to the same degree as arachidonic acid. Figure 4d shows that the two omega-3 PUFAs had a significant stimulatory effect in the neurite outgrowth assay, while the other indicated fatty acids were ineffective.

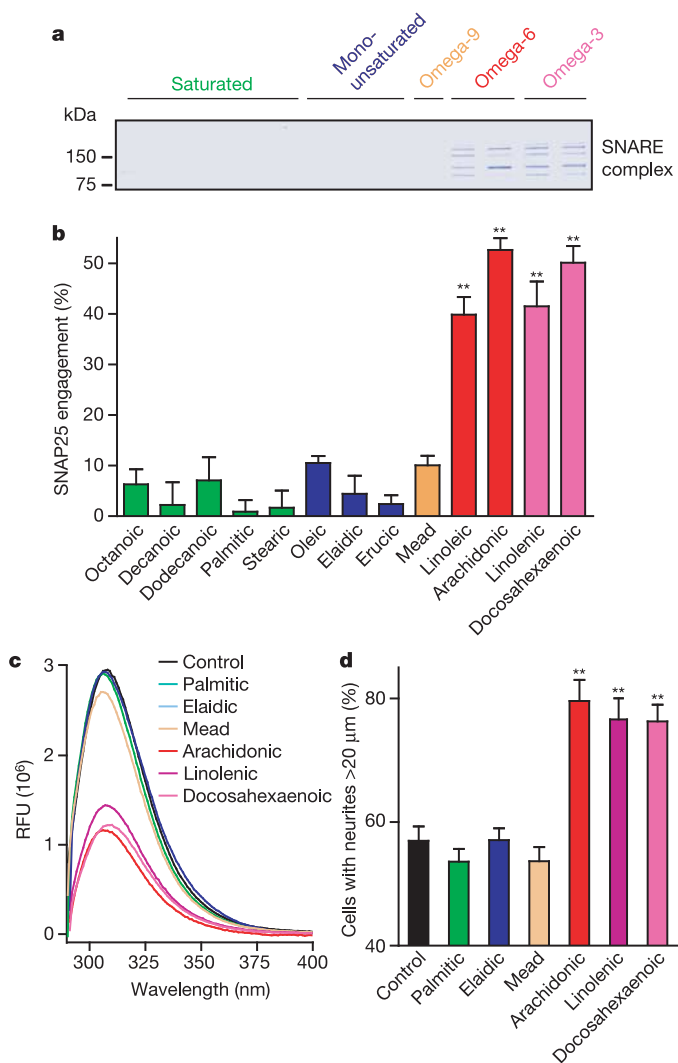


Figure 4 | Omega-3 and omega-6 PUFAs activate syntaxin 3 and stimulate neurite outgrowth. **a**, Screen of fatty acids for their ability to allow formation of the SDS-resistant SNARE complex reveals the specific importance of omega-3 and omega-6 PUFAs. A Coomassie-stained gel is shown. **b**, Quantification of the ability of fatty acids to stimulate the SNARE assembly, as measured by the engagement of SNAP25. **c**, Intrinsic protein fluorescence of syntaxin 3 is sensitive to omega-3 and omega-6 PUFAs (excitation wavelength 280 nm). **d**, Linolenic and docosahexaenoic acids are as active in the PC12 cell neurite outgrowth assay as arachidonic acid. Fatty acids incapable of syntaxin 3 activation fail to stimulate NGF-induced neurite outgrowth. The percentage of PC12 cells carrying neurites longer than 20 μm is shown. Fatty acids were tested at 200 μM concentration. Error bars are s.e.m.; $n = 4$; **, $P < 0.01$.

PUFAs, namely omega-3 and omega-6 fatty acids, are the subject of much scientific interest because of their important roles in human health, including neuronal development^{19,27}. As SNARE proteins act as gatekeepers for the addition of transport vesicle material to the plasma membrane, it is conceivable that some of the metabolic pathways activated during cell growth converge on SNARE-mediated function. Here we have demonstrated that syntaxin 3 is essential for neurite outgrowth, and is intrinsically inhibited—most likely through its own amino-terminal domain^{8,9}—even in the absence of Munc18 (also known as nSec1 or rbSec1) (Fig. 3). The ability of syntaxin 3 to partner with other SNAREs strictly requires omega-3 and omega-6 PUFAs, which act as if to ‘oil’ the plasma membrane fusion machinery. It is likely that PUFAs change the alpha-helical syntaxin structure to expose the SNARE motif for immediate SNAP25 engagement. Interestingly, most fatty-acid-binding proteins have a compact alpha-helical structure with hydrophobic crevices between amphipathic helices acting as binding sites for fatty acids²⁸. Syntaxin 3, being an integral plasma membrane protein, is ideally positioned to sense local transient increases in PUFA concentrations, and the threshold concentration required may ensure that syntaxin activation occurs only in the immediate vicinity of PLA2 action. The specific relevance of the omega-3 and omega-6 PUFAs for brain development^{19,27}, neuronal regeneration²⁹ and neurite outgrowth correlates closely with their ability to activate syntaxin 3 in fully defined reactions (Fig. 4). Together, our data demonstrate that syntaxin 3 is a single molecule effector for omega-3 and omega-6 PUFAs, raising the possibility of using this protein for the identification of new, potent compounds that can enhance neuronal growth. The ubiquitous nature of syntaxin 3 (ref. 14) indicates that the molecular mechanism for cell membrane expansion identified in this study could underlie cellular growth throughout the body.

METHODS

Plasmids and antibodies. Plasmids encoding GST fusions of rat SNAP25 (full length), syntaxin 3 (amino acids 1–261) and synaptobrevin 2 (amino acids 1–96) were previously described¹⁴. Mouse anti-syntaxin 1, mouse anti-βIII-tubulin and rabbit anti-syntaxin 3 were from Sigma. Rabbit anti-SNAP23 serum was from Synaptic Systems. Affinity-purified mouse anti-syntaxin 3, produced in-house against the GST-STX3 fusion, recognizes the syntaxin 3 isoform only. Mouse anti-SNAP25 was from Sternberger Monoclonals. Rabbit anti-SNAP25 has been previously described¹⁴.

Cell culture, neurite outgrowth and immunoprecipitation assays. PC12 cells and hippocampal neurons were cultured as described previously^{14,30}. For differentiation, PC12 cells were cultured in DMEM containing 2% horse serum, in the presence of 100 ng ml⁻¹ NGF 2.5-S subunit (Sigma). PC12 cells were treated with 10 μg ml⁻¹ PLA2 (ICN Biomedicals) for the duration of the NGF incubation. Fatty acids (Sigma) were dissolved in methanol, and methanol only was used in all control reactions. Cells were incubated with fatty acids for 30 min, washed three times with the culture medium, and grown in the presence of NGF for 48 h. Cell survival was quantified by a Trypan-blue exclusion assay. The siRNAs were purchased from Ambion, and were introduced into cells using the siPort Amine Reagent (Ambion) according to the manufacturer's instructions. Cells were immunostained as described previously¹⁴. For quantification of neurite outgrowth, PC12 cells were fixed and immunostained for βIII-tubulin. Images were collected using an ISIS-3 CCD camera (Photonic Science) and analysed with Scion Image software. Two hundred randomly chosen cells were analysed for each treatment condition in four independent experiments. Neurites were defined as thin processes longer than 5 μm when measured from the cell body to the end of the process. Immunoprecipitations were performed using Sepharose beads with covalently attached anti-SNAP25 antibody, or protein-G agarose beads carrying mouse anti-syntaxin 3 antibody as described previously¹⁴. Western immunoblotting was performed using the West Dura enhanced chemiluminescence kit (Pierce).

Protein reactions, circular dichroism and fluorescence assays. All reactions were performed in buffer A (20 mM HEPES, 100 mM NaCl, pH 7.3) for 30 min at 22 °C. Recombinant SNARE proteins were purified as previously described¹⁴. Size-exclusion chromatography was performed in buffer A on a Superdex-200 column (GE Healthcare). Binding of SNAP25 to GST-STX3, immobilized on Sepharose beads, was as previously described but in the absence of detergents, which, at high millimolar concentrations, can ‘open’ syntaxins^{12,14}. Formation of

the SDS-resistant SNARE complex was quantified as the difference in intensity of monomeric SNAP25, in the absence and presence of fatty acids, because the ternary complex itself has a complex migration pattern. Far-ultraviolet CD spectra were obtained by averaging five scans with a step size of 1 nm on a Jasco J-810 CD spectrometer. CD measurements were performed in a 0.2-mm path quartz cuvette (Hellma). For intrinsic fluorescence measurements, proteins in buffer A, with or without fatty acids, were added to a 1-cm path quartz cuvette (Hellma) and autofluorescence spectra collected on a Jobin-Yvon SPEX Fluoromax II fluorometer with excitation at 280 nm, integration time 0.1 s, excitation and emission slits of 5 nm. Note that syntaxin 3, unlike SNAP25, lacks tryptophans and exhibits only tyrosine autofluorescence. The ANS-based competition assay for fatty acid binding was performed as previously described²⁴.

Statistics. Band intensities in all biochemical experiments were quantified using a 12-bit CCD camera and analysed with Quantity One software (within the ChemiDoc XRS system, BioRad). Statistical analysis was performed using Prism4 (GraphPad). Multiple comparisons of means were performed using one-way analysis of variance (ANOVA) and Bonferroni tests.

Received 30 November; accepted 21 December 2005.

1. Futerman, A. H. & Banker, G. A. The economics of neurite outgrowth: the addition of new membrane to growing axons. *Trends Neurosci.* **19**, 144–149 (1996).
2. Negre-Aminou, P., Nemenoff, R. A., Wood, M. R., de la Houssaye, B. A. & Pfenninger, K. H. Characterization of phospholipase A2 activity enriched in the nerve growth cone. *J. Neurochem.* **67**, 2599–2608 (1996).
3. Hornfelt, M., Ekstrom, P. A. & Edstrom, A. Involvement of axonal phospholipase A2 activity in the outgrowth of adult mouse sensory axons *in vitro*. *Neuroscience* **91**, 1539–1547 (1999).
4. Kelly, R. B. Deconstructing membrane traffic. *Trends Cell Biol.* **9**, M29–M33 (1999).
5. Sofroniew, M. V., Howe, C. L. & Mobley, W. C. Nerve growth factor signaling, neuroprotection, and neural repair. *Annu. Rev. Neurosci.* **24**, 1217–1281 (2001).
6. Greene, L. A. & Tischler, A. S. Establishment of a noradrenergic clonal line of rat adrenal pheochromocytoma cells which respond to nerve growth factor. *Proc. Natl Acad. Sci. USA* **73**, 2424–2428 (1976).
7. Zheng, W. H., Fink, D. W. Jr & Guroff, G. Role of protein kinase C α in nerve growth factor-induced arachidonic acid release from PC12 cells. *J. Neurochem.* **66**, 1868–1875 (1996).
8. Rizo, J. & Sudhof, T. C. SNAREs and Munc18 in synaptic vesicle fusion. *Nature Rev. Neurosci.* **3**, 641–653 (2002).
9. Ungar, D. & Hughson, F. M. SNARE protein structure and function. *Annu. Rev. Cell Dev. Biol.* **19**, 493–517 (2003).
10. Wenk, M. R. & De Camilli, P. Protein–lipid interactions and phosphoinositide metabolism in membrane traffic: insights from vesicle recycling in nerve terminals. *Proc. Natl Acad. Sci. USA* **101**, 8262–8269 (2004).
11. Rohrbough, J. & Broadie, K. Lipid regulation of the synaptic vesicle cycle. *Nature Rev. Neurosci.* **6**, 139–150 (2005).
12. Rickman, C. & Davletov, B. Arachidonic acid allows SNARE complex formation in the presence of Munc18. *Chem. Biol.* **12**, 545–553 (2005).
13. Tang, B. L. Protein trafficking mechanisms associated with neurite outgrowth and polarized sorting in neurons. *J. Neurochem.* **79**, 923–930 (2001).
14. Bajohrs, M., Darios, F., Peak-Chew, S. Y. & Davletov, B. Promiscuous interaction of SNAP-25 with all plasma membrane syntaxins in a neuroendocrine cell. *Biochem. J.* **392**, 283–289 (2005).
15. Schiavo, G., Shone, C. C., Bennett, M. K., Scheller, R. H. & Montecucco, C. Botulinum neurotoxin type C cleaves a single Lys–Ala bond within the carboxyl-terminal region of syntaxins. *J. Biol. Chem.* **270**, 10566–10570 (1995).
16. Igarashi, M. *et al.* Growth cone collapse and inhibition of neurite growth by Botulinum neurotoxin C1: a t-SNARE is involved in axonal growth. *J. Cell Biol.* **134**, 205–215 (1996).
17. Osen-Sand, A. *et al.* Inhibition of axonal growth by SNAP-25 antisense oligonucleotides *in vitro* and *in vivo*. *Nature* **364**, 445–448 (1993).
18. Farkas, T. *et al.* Docosahexaenoic acid-containing phospholipid molecular species in brains of vertebrates. *Proc. Natl Acad. Sci. USA* **97**, 6362–6366 (2000).
19. Wainwright, P. E. Dietary essential fatty acids and brain function: a developmental perspective on mechanisms. *Proc. Nutr. Soc.* **61**, 61–69 (2002).
20. Brash, A. R. Arachidonic acid as a bioactive molecule. *J. Clin. Invest.* **107**, 1339–1345 (2001).
21. Negre-Aminou, P. & Pfenninger, K. H. Arachidonic acid turnover and phospholipase A2 activity in neuronal growth cones. *J. Neurochem.* **60**, 1126–1136 (1993).
22. Fasshauer, D., Bruns, D., Shen, B., Jahn, R. & Brunger, A. T. A structural change occurs upon binding of syntaxin to SNAP-25. *J. Biol. Chem.* **272**, 4582–4590 (1997).
23. Lakowicz, J. R. *Principles of Fluorescence Spectroscopy* (Kluwer Academic/Plenum, New York, 1999).
24. Kane, C. D. & Bernlohr, D. A. A simple assay for intracellular lipid-binding proteins using displacement of 1-anilinonaphthalene 8-sulfonic acid. *Anal. Biochem.* **233**, 197–204 (1996).
25. Fasshauer, D., Antonin, W., Margittai, M., Pabst, S. & Jahn, R. Mixed and non-cognate SNARE complexes. Characterization of assembly and biophysical properties. *J. Biol. Chem.* **274**, 15440–15446 (1999).
26. Galli, T. *et al.* A novel tetanus neurotoxin-insensitive vesicle-associated membrane protein in SNARE complexes of the apical plasma membrane of epithelial cells. *Mol. Biol. Cell* **9**, 1437–1448 (1998).
27. Marszalek, J. R. & Lodish, H. F. Docosahexaenoic acid, fatty acid-interacting proteins, and neuronal function: breastmilk and fish are good for you. *Annu. Rev. Cell Dev. Biol.* **21**, 633–657 (2005).
28. Hamilton, J. A. Fatty acid interactions with proteins: what X-ray crystal and NMR solution structures tell us. *Prog. Lipid Res.* **43**, 177–199 (2004).
29. Nakamura, S. Involvement of phospholipase A2 in axonal regeneration of brain noradrenergic neurones. *Neuroreport* **4**, 371–374 (1993).
30. Bading, H. & Greenberg, M. E. Stimulation of protein tyrosine phosphorylation by NMDA receptor activation. *Science* **253**, 912–914 (1991).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Drevier for neuronal cultures. F.D. was supported by an EMBO long-term fellowship.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to B.D. (email@bazbek.com).

LETTERS

Nck adaptor proteins link nephrin to the actin cytoskeleton of kidney podocytes

Nina Jones¹, Ivan M. Blasutig^{1,2*}, Vera Eremina^{1,3*}, Julie M. Ruston¹, Friedhelm Blatt¹, Hongping Li⁴, Haiming Huang⁵, Louise Larose⁴, Shawn S.-C. Li⁵, Tomoko Takano⁴, Susan E. Quaggin^{1,3,6} & Tony Pawson^{1,2}

The glomerular filtration barrier in the kidney is formed in part by a specialized intercellular junction known as the slit diaphragm, which connects adjacent actin-based foot processes of kidney epithelial cells (podocytes)¹. Mutations affecting a number of slit diaphragm proteins, including nephrin (encoded by *NPHS1*)², lead to renal disease owing to disruption of the filtration barrier and rearrangement of the actin cytoskeleton³, although the molecular basis for this is unclear. Here we show that nephrin selectively binds the Src homology 2 (SH2)/SH3 domain-containing Nck adaptor proteins⁴, which in turn control the podocyte cytoskeleton *in vivo*. The cytoplasmic tail of nephrin has multiple YDxV sites that form preferred binding motifs for the Nck SH2 domain once phosphorylated by Src-family kinases. We show that this Nck–nephrin interaction is required for nephrin-dependent actin reorganization. Selective deletion of Nck from podocytes of transgenic mice results in defects in the formation of foot processes and in congenital nephrotic syndrome. Together, these findings identify a physiological signalling pathway in which nephrin is linked through phosphotyrosine-based interactions to Nck adaptors, and thus to the underlying actin cytoskeleton in podocytes. Simple and widely expressed SH2/SH3 adaptor proteins can therefore direct the formation of a specialized cellular morphology *in vivo*.

Nephrin is a 180-kDa transmembrane protein of the immunoglobulin superfamily that functions as an adhesion molecule and a structural component of the slit diaphragm^{2,5}. Nephrin also serves as a signalling scaffold by recruiting proteins to the cytoplasmic face of podocyte foot processes⁶, including CD2AP and podocin, which are necessary for the maintenance of normal podocyte structure^{7,8}, as well as IQGAP1, MAGIs, spectrins and α -actinin⁹. Consistent with a role in signalling, nephrin is tyrosine-phosphorylated and associates with Src-family kinases *in vivo*^{10,11}, which phosphorylate the cytoplasmic tail of nephrin upon clustering *in vitro*¹². Genetic evidence supports a role for Src-family kinases in the maintenance of glomerular function, as mice deficient in Fyn or Fyn and Yes develop foot process effacement (loss) and proteinuria^{10,13}. However, the physiological significance of phosphotyrosine-based nephrin signalling at the slit diaphragm, and its connection to the cytoskeleton, remain poorly understood.

The cytoplasmic region of nephrin has a series of conserved tyrosine-based motifs¹⁴ (Fig. 1a), which upon phosphorylation can serve as binding sites for the SH2 domains of cytoplasmic targets such as phosphatidylinositol-3-OH kinase¹⁵ and Fyn¹⁰. Notably, three tyrosine residues in the intracellular regions of human and mouse nephrin lie within potential binding sites for the SH2 domain

of Nck adaptor proteins¹⁴. Nck1 (also Nck α) and Nck2 (also Nck β or Grb4) comprise a family of adaptor proteins (collectively 'Nck') with three SH3 domains and a carboxy-terminal SH2 domain¹⁶ (Fig. 1b). We have previously found that the SH2 domain of Nck binds with high affinity to a phosphorylated motif in the cytoplasmic tail of the Tir protein of enteropathogenic *E. coli* with the core sequence YDEV (Fig. 1c)¹⁷. Previous analysis of the binding properties of the Nck1 SH2 domain using a soluble degenerate peptide library has yielded the consensus motif phospho-(p)Tyr-Asp-Glu-Pro>Asp>Val. We have now probed an oriented phosphopeptide array library with the Nck1 SH2 domain, and find a strong selection for Asp in the +1 position relative to pTyr and for Val at +3 (Fig. 1d). Consistent with the data from enteropathogenic *E. coli* Tir and the oriented phosphopeptide array library, two of the potential Nck SH2-binding sites in nephrin are YDEV motifs, and the third is YDQV. Nck interacts through its SH3 domains with various effector proteins that regulate cytoskeletal organization, including N-WASp, which stimulates the Arp2/3 complex to initiate actin polymerization, and the protein kinase Pak^{4,18,19}. Recruitment of Nck to phosphorylated YDxV sites on nephrin could therefore directly control the cytoskeletal architecture of podocytes (Fig. 1b).

To determine whether the SH2 domain of Nck adaptors can associate with nephrin, an immobilized glutathione S-transferase (GST)–Nck1 SH2 fusion protein was incubated with lysates from nephrin-transfected HEK 293T cells. Transiently expressed nephrin was not tyrosine-phosphorylated at detectable levels. However, treatment of these cells with the tyrosine phosphatase inhibitor pervanadate or co-expression of active Src kinase induced tyrosine phosphorylation of nephrin and its association with the Nck1 SH2 domain (Fig. 2a). To pursue these findings in the context of full-length Nck, we next analysed the association of transfected human nephrin with endogenous Nck adaptors in HEK 293T cells. Tyrosine-phosphorylated nephrin selectively co-precipitated with both Nck1 and Nck2 (Fig. 2b and data not shown). We also examined whether the interaction between nephrin and Nck could be detected in glomeruli from adult rats. Nephrin from glomerular lysates was tyrosine phosphorylated, and bound *in vitro* to the Nck1 SH2 domain (Fig. 2c and data not shown). Furthermore, we observed co-immunoprecipitation of phosphorylated nephrin and Nck from glomeruli (Fig. 2c). These experiments indicate that Nck binds through its SH2 domain to tyrosine-phosphorylated nephrin both in transfected cells and *in vivo*.

To address whether the YDxV sites in nephrin are required for association with Nck proteins, we introduced single, double or triple

¹Samuel Lunenfeld Research Institute, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada. ²Department of Molecular and Medical Genetics, 4388 Medical Sciences Building, 1 King's College Circle, University of Toronto, Toronto, Ontario M5S 1A8, Canada. ³Institute of Medical Science, 7213 Medical Sciences Building, 1 King's College Circle, University of Toronto, Toronto, Ontario M5S 1A8, Canada. ⁴Department of Medicine, McGill University Health Centre, 3775 University Street, Montreal, Québec H3A 2B4, Canada. ⁵Department of Biochemistry, University of Western Ontario, London, Ontario N6A 5C1, Canada. ⁶Division of Nephrology, St. Michael's Hospital, 30 Bond Street, Toronto, Ontario M5B 1W8, Canada.

*These authors contributed equally to this work.

tyrosine (Y) to phenylalanine (F) substitutions into full-length human nephrin, at amino acid positions 1176, 1193 and 1217. Transient expression of these mutants in HEK 293T cells with active Src indicated that nephrin tyrosine phosphorylation was decreased in the double and triple mutants, suggesting that these sites might be substrates for Src-family kinases in cells (Fig. 2d). Consistent with this possibility, each nephrin YDxV motif is preceded by a leucine or isoleucine residue, which are the preferred amino acids at the -1 position for phosphorylation by Src-family kinases. Indeed, an *in vitro* kinase assay confirmed that the YDxV motifs on nephrin were preferentially phosphorylated by Fyn and Src (Fig. 1e), as previously seen with rat nephrin¹¹. We then used lysates from HEK 293T cells transfected with nephrin mutants in co-immunoprecipitation experiments. Mutant forms of human nephrin with single Y-to-F substitutions all associated with Nck (Fig. 2d). However, the interaction of Nck with nephrin was selectively reduced by combined Y1193F/Y1217F substitutions, and was virtually undetectable for the Y1176F/Y1217F nephrin mutant (Fig. 2d). No binding of Nck to nephrin was detected when all three YDxV tyrosines were mutated to phenylalanine (Y3F) (Fig. 2d). The involvement of multiple YDxV motifs in Nck binding was also observed in rat and mouse nephrin (data not shown). Furthermore, the SH2 domain of Nck1 selectively associated with phosphopeptides representing each of the YDxV

motifs on nephrin (Fig. 1e), indicating that nephrin has multiple docking sites for Nck.

The *in vivo* interaction with nephrin implies that Nck might have an important function in the kidney slit diaphragm. Consistent with this possibility, both Nck1 and Nck2 are expressed in podocytes (Supplementary Fig. 1a, b, e). However, mice lacking either *Nck1* or *Nck2* are viable²⁰ and show no apparent renal defects (Supplementary Fig. 1c, d), suggesting that Nck1 and Nck2 might have overlapping functions. Indeed, mice homozygous for both *Nck1*- and *Nck2*-null alleles die at embryonic day 9.5 (ref. 20), precluding analysis of Nck function in podocytes of double-null animals. To circumvent this lethality, we used a conditional *loxP*-flanked (floxed) allele of *Nck2* (J. Fawcett, F.B. and T.P., unpublished results) to selectively ablate *Nck2* expression in podocytes of *Nck1*-null (*Nck1*^{-/-}) animals by expressing Cre recombinase under the control of the podocyte-specific podocin promoter (Supplementary Fig. 1f). Mice homozygous for both the *Nck1*-null allele and a *Nck2* floxed allele, and carrying one copy of a podocin-Cre transgene (podocin-Cre^{+/+}; *Nck1*^{-/-}, *Nck2*^{fl/fl}, hereafter designated *Nck* mutant mice) were born at the expected mendelian frequency. However, by four days of age, *Nck* mutant pups were smaller than their littermates, and this growth retardation became more apparent by three weeks of age (Supplementary Fig. 2a). Urinalysis from *Nck* mutant mice showed

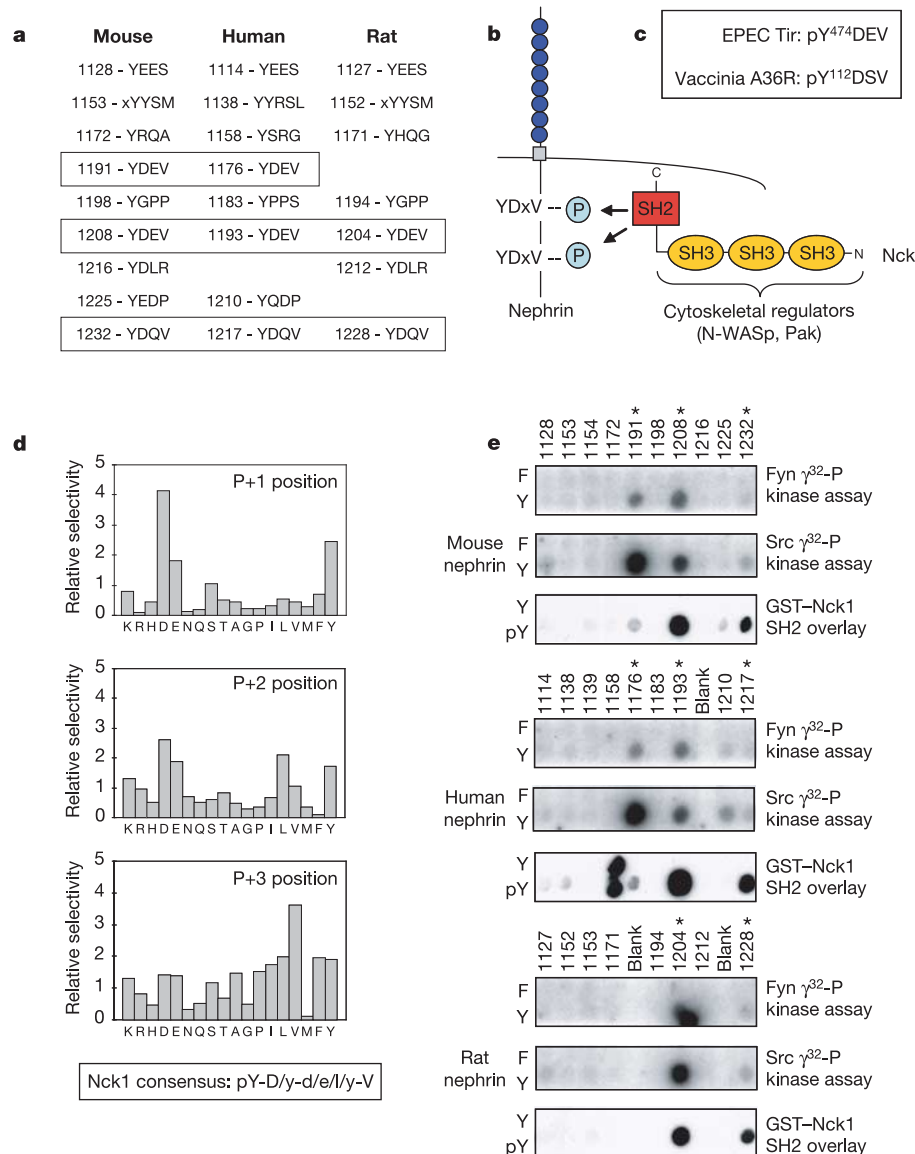


Figure 1 | The kidney slit diaphragm protein nephrin has multiple YDxV motifs that bind the Nck SH2 domain. **a**, Tyrosine residues and conserved YDxV motifs (boxed) in mouse, human and rat nephrin. Numbers indicate the amino acid position of the (first) tyrosine residue. **b**, Recruitment of Nck to phosphorylated YDxV sites on nephrin and downstream activation of N-WASP and Pak. **c**, Phosphorylated YDxV motifs shown to recruit Nck in enteropathogenic *E. coli* (EPEC) Tir and vaccinia A36R. **d**, Oriented phosphopeptide array libraries yield a consensus binding motif for the Nck1 SH2 domain of pY-D/y-d/e/l/y-V (residues under strong selection in upper case; residues under weak selection in lower case). **e**, Spot peptide arrays (centred around F, Y or phospho- (p)Y) were subjected to an *in vitro* kinase assay using recombinant Fyn or Src, or incubated with purified GST-Nck1 SH2 domain or GST alone. Binding to human Y1158 was also seen with GST alone (not shown). Asterisks indicate tyrosine residues within YDxV motifs.

nephrotic-range proteinuria, and SDS–polyacrylamide gel electrophoresis confirmed large amounts of albumin in the urine (Supplementary Fig. 2b). Histological examination of kidneys also revealed proteinaceous material within the tubules of mutant mice (Supplementary Fig. 2c, d). Most individual glomeruli remained relatively normal in the first few weeks of life in *Nck* mutant mice, as assessed by light microscopy (Fig. 3a–d). However, by 3.5 weeks of age, a range of glomerular defects were detected, including focal sclerosis (Fig. 3e). In some cases, much of the glomerulus had been replaced with sclerotic tissue, consistent with end-stage kidney disease (Fig. 3f). Expression of molecular markers of glomerular development (Wilm's tumour-1 (WT-1) and nephrin) in the glomeruli of *Nck* mutant mice was comparable to littermate controls at one week of age (Supplementary Fig. 3a–d) but was reduced by 3.5 weeks of age in a pattern consistent with segmental changes in podocyte distribution (Supplementary Fig. 3e–h), indicating a loss of podocytes.

Although the glomeruli in neonatal mutant mice did not seem morphologically abnormal by light microscopy examination (Fig. 3a–d), the presence of albuminuria indicated that the capillary filtration barrier might be damaged. This barrier normally consists of an inner fenestrated endothelium that is separated by a glomerular basement membrane from podocytes, with their characteristic foot processes (Fig. 3g). Electron micrographic examination of kidneys from four-day-old pups showed complete fusion of foot processes around the capillary loops of fully differentiated glomeruli in *Nck* mutant animals, compared to the regular appearance of these structures in control littermates (Fig. 3h, i). To determine whether foot processes can actually form in mutant mice or whether they develop and then degenerate, we performed similar ultrastructural

analysis on kidney sections from embryos at day 16.5 to visualize podocytes that are beginning to differentiate and establish foot processes. In glomeruli from both mutant and control embryos, endothelial cell fenestrations were present and podocytes spread normally around the capillary loops (Fig. 3j, k). However, differentiated foot processes were notably absent at this stage of development in *Nck* mutant embryos compared to controls, consistent with the hypothesis that foot processes fail to form in *Nck* mutant animals.

The effacement of podocyte foot processes typically arises owing to perturbations in the actin cytoskeleton³. *Nck* adaptor proteins have been implicated in connecting phosphotyrosine signals to the actin cytoskeleton through SH3-mediated interactions with molecular effectors. Indeed, clustering of *Nck* SH3 domains at the plasma membrane is sufficient to induce localized actin polymerization²¹. We therefore investigated whether recruitment of *Nck* signalling complexes to the plasma membrane in response to nephrin clustering can stimulate reorganization of the actin cytoskeleton. To specifically activate intracellular nephrin signalling, we fused the intracellular (IC) domains of either wild-type (WT) nephrin or the *Nck*-binding mutant (Y3F) to the extracellular and transmembrane domains of the human immunoglobulin F_c receptors CD16 and CD7, respectively, to generate CD16/7–nephrin(WT)^{IC} and CD16/7–nephrin(Y3F)^{IC} chimaeras. These chimaeric proteins, which also included C-terminal green fluorescent protein (GFP), were expressed in murine embryonic fibroblasts (MEFs) lacking endogenous *Nck1* and *Nck2*²⁰. Clustering of these fusion proteins by treatment with anti-CD16 antibody resulted in markedly enhanced tyrosine phosphorylation of CD16/7–nephrin(WT)^{IC} and increased binding of a co-transfected Flag-tagged version of *Nck2* or its target N-WASP (Fig. 4a and Supplementary Fig. 4). In contrast, neither nephrin

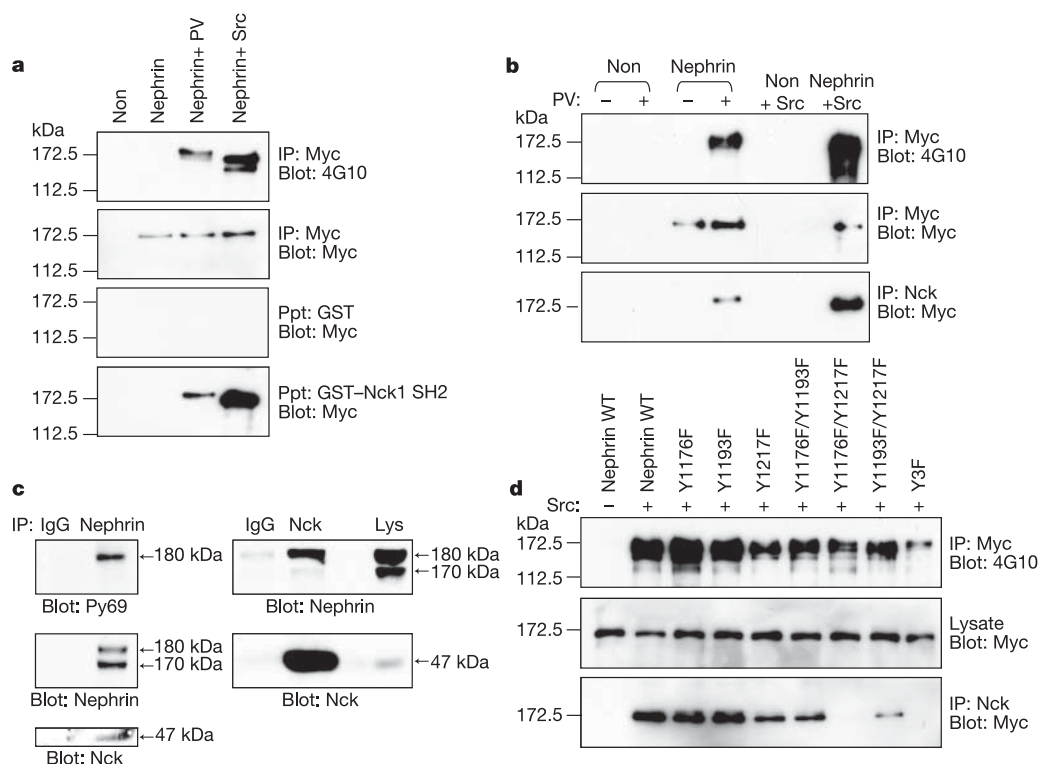


Figure 2 | Identification of a phosphotyrosine-dependent interaction between nephrin and the *Nck* SH2/SH3 adaptor. **a**, Myc-tagged nephrin is tyrosine-phosphorylated in HEK 293T cells stimulated with pervanadate (PV) or co-transfected with active Src, and can be precipitated (Ppt) by a GST–*Nck1* SH2 domain. The 4G10 antibody is used to detect phosphotyrosine. Non, untransfected cells. **b**, Endogenous *Nck* associates exclusively with tyrosine-phosphorylated nephrin in parallel HEK 293T transfectants. **c**, In adult rat glomeruli, endogenous nephrin appears as a

doublet of 180 and 170 kDa, and the upper band (180 kDa) is tyrosine-phosphorylated. Tyrosine-phosphorylated nephrin co-immunoprecipitates specifically with *Nck* (47 kDa) *in vivo*. Lys, total lysate. Py69 is another antibody that recognizes phosphotyrosine. **d**, Tyrosine-to-phenylalanine substitutions in nephrin variably reduce nephrin tyrosine phosphorylation and association with *Nck*. Interaction of nephrin with *Nck* is completely undetectable with the Y3F mutant. IP, immunoprecipitate; WT, wild type.

phosphorylation nor Nck2/N-WASp recruitment were detectably increased upon clustering of the CD16/7–nephrin(Y3F)^{IC} mutant (Fig. 4a and Supplementary Fig. 4). Clustering-induced tyrosine phosphorylation of nephrin seemed dependent on the action of Src-family kinases, as inhibitors specific for these kinases decreased this response, and phosphorylation was not observed in a MEF line lacking Src, Yes and Fyn (Supplementary Fig. 5). We confirmed by immunofluorescence microscopy that GFP aggregates corresponding to CD16/7–nephrin^{IC} proteins were the result of CD16 clustering by using a fluorescently labelled secondary antibody against CD16 in live cells (Fig. 4b).

To examine whether nephrin signalling influences actin organization, cells expressing CD16/7–nephrin^{IC} were fixed and stained with phalloidin to visualize F-actin after antibody-induced clustering. For this purpose, we used either MEFs lacking both *Nck1* and *Nck2*, or cells in which *Nck2* expression had been restored, allowing us to test directly whether Nck adaptors are involved in coupling clustered nephrin to actin polymerization. In the absence of Nck, expression and clustering of CD16/7–nephrin(WT)^{IC} had no observable effect on actin organization (Fig. 4c, row 3). In contrast, co-expression of *Nck2* in these same cells followed by clustering of CD16/7–nephrin(WT)^{IC} using an anti-CD16 antibody resulted in striking rearrangement of the actin cytoskeleton, notably in the form of polymerized actin ‘comet tails’ (Fig. 4c, row 4 and Fig. 4d). Before CD16 crosslinking, transfected *Nck2* did not colocalize with CD16/7–nephrin(WT)^{IC} but remained diffusely distributed in a cytoplasmic pattern, as observed when *Nck2* was expressed alone (Fig. 4c, rows 1 and 2). However, upon CD16 crosslinking, *Nck2* became colocalized with clusters of CD16/7–nephrin(WT)^{IC}, which

visibly capped many of the tails of polymerized actin induced by nephrin aggregation (the white ‘heads’ in Fig. 4c, row 4 and Fig. 4d). Notably, when the CD16/7–nephrin(Y3F)^{IC} mutant was co-expressed with *Nck2* in *Nck*-deficient MEFs, *Nck2* was not relocalized upon anti-CD16 treatment, nor was there any overt effect on actin polymerization, despite the fact that the Y3F mutant was aggregated upon crosslinking with anti-CD16 (Fig. 4c, row 5 and Fig. 4e). Similar results were obtained upon clustering of full-length nephrin (data not shown). Taken together, these results show that Nck adaptor proteins are required to link phosphorylated YDxV sites on activated nephrin to reorganization of the actin cytoskeleton.

Congenital foot process effacement due to loss of Nck in podocytes resembles that observed in mice deficient in nephrin^{22,23}, as well as other components of the slit diaphragm such as *Neph1*²⁴, *FAT1*²⁵ and α_3 integrin²⁶. Although it is possible that Nck also acts downstream of other transmembrane podocyte proteins such as $\alpha_3\beta_1$ integrin, these molecules lack direct Nck-binding motifs. The results presented here argue that direct interaction of Nck with nephrin drives the elaboration of actin-based foot processes *in vivo*. The importance of actin remodelling in podocyte function is supported by genetic evidence in patients with adult-onset focal segmental glomerulosclerosis, who carry mutations in the actin filament crosslinking protein α -actinin-4 (*ACTN4*)²⁷. Through interaction with α -actinin-4 and the Arp2/3 complex, the cytoplasmic adaptor CD2AP has also been proposed to anchor nephrin to the actin cytoskeleton²⁸. Although mice lacking CD2AP or α -actinin-4 develop nephrotic syndrome, foot process formation is initiated in these mice^{7,29}, in contrast to *Nck*-deficient animals, suggesting that this complex might primarily be required for the maintenance of podocyte foot processes. The roles of Nck and

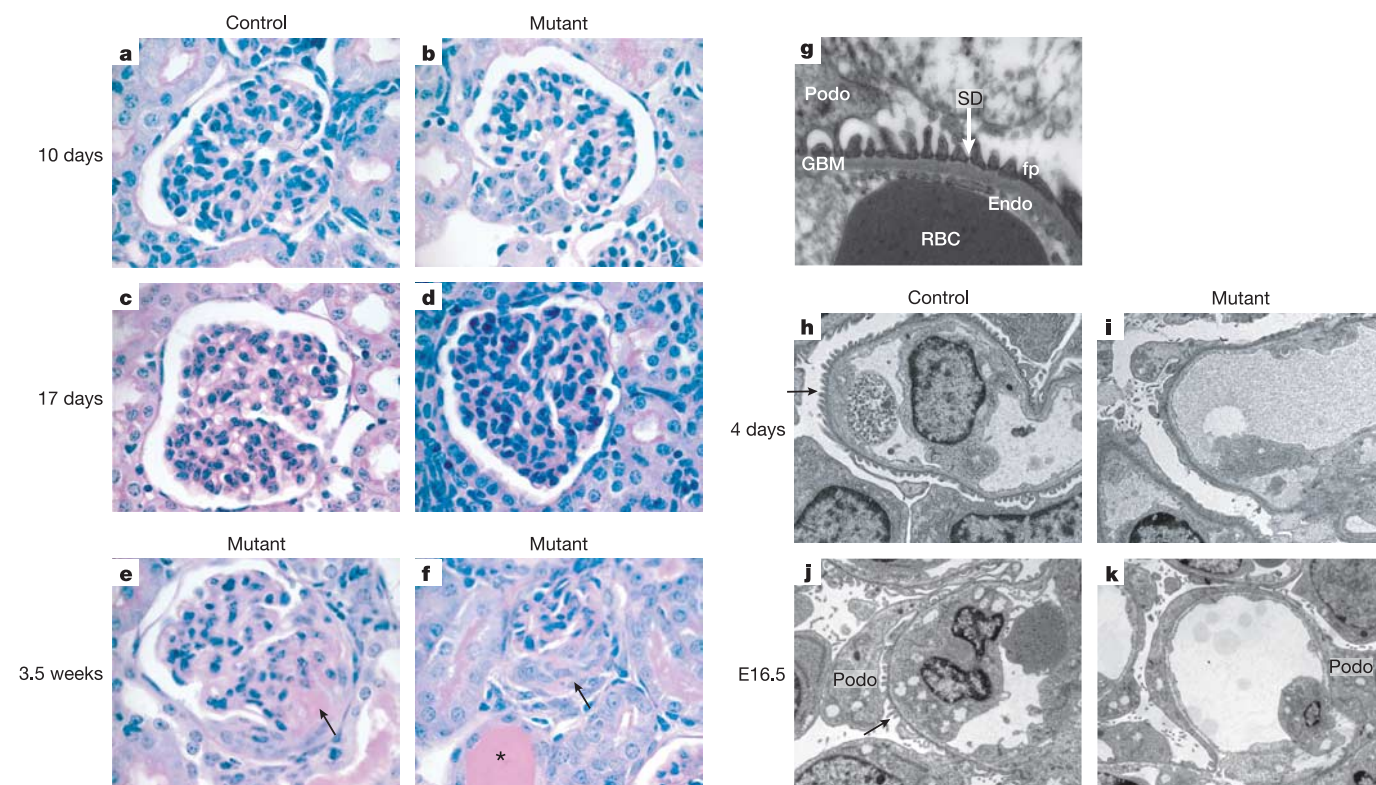


Figure 3 | Glomerulosclerosis and foot process fusion in mice lacking Nck expression in podocytes. **a–f**, Periodic Acid Schiff (PAS) staining of kidney sections reveals that *Nck* mutant glomeruli (**b**, **d**) are comparable to *Cre*^{-/-} controls (**a**, **c**) at 10 and 17 days of age. However, after 3.5 weeks, both focal (**e**) and widespread (**f**) sclerosis (arrows) can be observed in mutant glomeruli. Asterisk indicates kidney tubule dilation and protein deposits. **g**, Electron micrograph of the normal glomerular filtration barrier. The slit diaphragm (SD) formed between adjacent foot processes (fp) is indicated.

Endo, endothelium; GBM, glomerular basement membrane; Podo, podocytes; RBC, red blood cell. **h**, **i**, Distinct foot processes (arrow) normally seen at 4 days of age (**h**) cannot be distinguished in a comparable capillary loop from a *Nck* mutant (**i**), but are instead flattened, with an absence of slit diaphragms. **j**, **k**, At embryonic day (E)16.5, podocyte foot processes (arrow) form around a developing capillary loop in control littermates (**j**), but remain spread around the loop in *Nck* mutants (**k**).

CD2AP in signalling to the podocyte cytoskeleton are therefore likely to be distinct, and it will now be of interest to explore whether Nck expression is also required in the mature glomerular filtration barrier.

Genetic data support the function of nephrin as a signalling platform, as deletion of the cytoplasmic region of nephrin in patients with the *Fin_{minor}* mutation results in a similar disease phenotype to complete loss of nephrin². This region of nephrin seems tailored to regulate the actin cytoskeleton via Nck through the presence of multiple YDxV motifs. Notably, such YDxV phosphorylation sites

are also present in the bacterial EPEC Tir protein and the A36R vaccinia virus polypeptide (Fig. 1c), both of which become phosphorylated by cellular kinases, bind Nck and trigger actin polymerization through recruitment of effectors such as N-WASP and Arp2/3^{17,30}. Our data indicate that these pathogens have mimicked the mechanism normally used by cell-surface proteins such as nephrin to reorganize the cytoskeleton, through the simple acquisition of YDxV sites. In addition to interacting with Nck, nephrin associates with several cytoplasmic proteins at the slit diaphragm⁶, which may

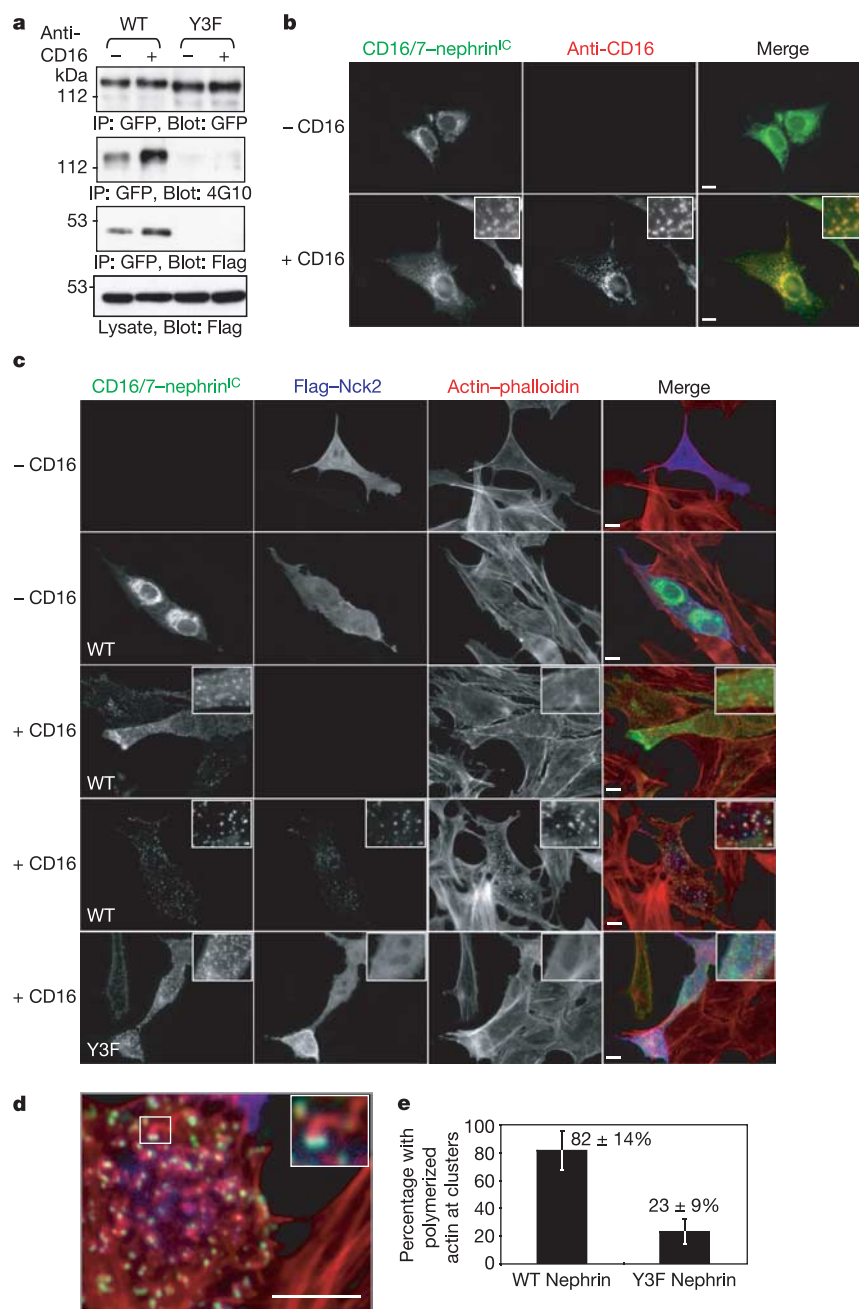


Figure 4 | Nck recruitment to phosphorylated nephrin facilitates localized actin reorganization. **a**, In *Nck1*^{-/-}, *Nck2*^{-/-} MEFs expressing GFP-tagged CD16/7-nephrin^{IC} chimaeric proteins with Flag-Nck2, wild-type (WT) nephrin but not Y3F nephrin becomes tyrosine-phosphorylated and binds Nck2 after anti-CD16 treatment. **b**, Parallel transfectants were imaged using fluorescence microscopy to demonstrate co-localization of cell-surface CD16 (red) and intracellular GFP (green). **c**, Relocalization of Nck2 (blue) and comet tails of actin (red) can be seen in stimulated MEFs co-expressing both CD16/7-nephrin(WT)^{IC} (green) and Nck2 (row 4) but

not in unstimulated (row 2) or singly transfected (rows 1, 3) cells, or in cells co-expressing CD16/7-nephrin(Y3F)^{IC} and Nck2 (row 5). The merged images in row 4 and panel **d** illustrate the overlap between CD16/7-nephrin(WT)^{IC} and Nck2, where the clusters appear as white 'heads' with red 'tails' of polymerized actin. Scale bars, 10 μ m.

e, Quantification of transfected *Nck*-null MEFs in rows 4 and 5 of **c**, showing polymerized actin at sites of nephrin clustering, expressed as a percentage of 100 cells counted at random in three independent experiments. Error bars show s.d.

modulate the cellular response to Nck signalling. Podocytes therefore present an intriguing system to explore how a widely expressed adaptor such as Nck interacts in a combinatorial fashion with more selectively expressed proteins to sculpt an elaborate actin-based cellular morphology *in vivo*.

METHODS

Plasmids. Myc-tagged human nephrin cDNA was provided by T. Huber. Tyrosine-to-phenylalanine substitutions were carried out using polymerase chain reaction (PCR)-based mutagenesis with Pfx Platinum polymerase (Invitrogen), and were confirmed by DNA sequencing. A plasmid containing the extracellular domain of CD16 and the transmembrane domain of CD7 was provided by G. Rivera and B. Mayer. The intracellular domain of human nephrin was PCR-amplified to generate an in-frame fusion with CD16/7 followed by a C-terminal EGFP epitope tag. Full-length human *NCK1* (NCBI accession number BC006403) and *NCK2* (BC000103) cDNAs (Open Biosystems) were PCR-amplified with a C-terminal Flag epitope tag and cloned into pcDNA3 (Invitrogen), or the SH2 domains of *NCK1* or *NCK2* were cloned into pGEX-4T-1 (Amersham Biosciences) to generate GST fusion proteins. Flag-N-WASP has been described previously¹⁷.

Cell culture. All cell lines were grown in DMEM medium supplemented with 10% fetal bovine serum (Hyclone). Transient transfections were performed using polyethylenimine for HEK 293T cells or Effectene reagent (Qiagen) for MEFs. Pervanadate (50 μ M) stimulation of HEK 293T cells was performed for 10 min at 37 °C. MEFs were pre-treated with kinase inhibitors as indicated for 2 h before stimulation with anti-CD16. Protein blotting techniques and spot peptide arrays are described in Supplementary Information.

CD16 crosslinking and cell imaging. Antibody-mediated crosslinking of CD16/7 fusion proteins has been described previously²¹. Briefly, 1.5×10^5 MEFs were seeded onto glass coverslips and transfected the next day. After ~18 h, cells were incubated with $1 \mu\text{g ml}^{-1}$ mouse monoclonal anti-CD16 (Sigma) for 15 min at 37 °C, washed once with DMEM, incubated with $1 \mu\text{g ml}^{-1}$ fluorescence-conjugated secondary antibody for 15 min and fixed in 4% paraformaldehyde. Staining for Flag-Nck2 was performed using a polyclonal anti-Flag antibody (Sigma) and actin was visualized using fluorescently labelled phalloidin (Molecular Probes). For stimulation of transfected cells in preparation for phosphotyrosine analysis, anti-CD16 antibody was added at a dilution of 1:200 in culture medium for 10 min at 37 °C.

Generation of podocin-Cre^{+/+}; Nck1^{-/-}, Nck2^{fl/fl} mutant mice. Mice carrying a floxed *Nck2* allele (*Nck2*^{fl/fl}) in which exon 1 is flanked by *loxP* sites to allow Cre-mediated excision (J. Fawcett, E.B. and T.P., unpublished results) and a *Nck1*^{-/-} allele²⁰ were crossed to transgenic mice expressing Cre recombinase under the control of the podocyte-specific podocin (*NPHS2*) promoter (podocin-Cre). Cre-positive heterozygous offspring were backcrossed to *Nck1*^{-/-}, *Nck2*^{fl/fl} mice to obtain podocin-Cre^{+/+}; *Nck1*^{-/-}, *Nck2*^{fl/fl} mutant mice. Generation of *Nck2*^{-/-} mice used for *lacZ* expression analysis has been described elsewhere²⁰. Animal husbandry was carried out in accordance with Canadian Council on Animal Care protocols.

Received 12 January; accepted 21 February 2006.

Published online 8 March 2006.

- Pavenstadt, H., Kriz, W. & Kretzler, M. Cell biology of the glomerular podocyte. *Physiol. Rev.* **83**, 253–307 (2003).
- Kestila, M. *et al.* Positionally cloned gene for a novel glomerular protein—nephrin—is mutated in congenital nephrotic syndrome. *Mol. Cell* **1**, 575–582 (1998).
- Oh, J., Reiser, J. & Mundel, P. Dynamic (re)organization of the podocyte actin cytoskeleton in the nephrotic syndrome. *Pediatr. Nephrol.* **19**, 130–137 (2004).
- Buday, L., Wunderlich, L. & Tamas, P. The Nck family of adapter proteins: regulators of actin cytoskeleton. *Cell. Signal.* **14**, 723–731 (2002).
- Ruotsalainen, V. *et al.* Nephrin is specifically located at the slit diaphragm of glomerular podocytes. *Proc. Natl Acad. Sci. USA* **96**, 7962–7967 (1999).
- Benzing, T. Signaling at the slit diaphragm. *J. Am. Soc. Nephrol.* **15**, 1382–1391 (2004).
- Shih, N. Y. *et al.* Congenital nephrotic syndrome in mice lacking CD2-associated protein. *Science* **286**, 312–315 (1999).
- Roselli, S. *et al.* Early glomerular filtration defect and severe renal disease in podocin-deficient mice. *Mol. Cell. Biol.* **24**, 550–560 (2004).
- Lehtonen, S. *et al.* Cell junction-associated proteins IQGAP1, MAGI-2, CASK, spectrins, and α -actinin are components of the nephrin multiprotein complex. *Proc. Natl Acad. Sci. USA* **102**, 9814–9819 (2005).

- Verma, R. *et al.* Fyn binds to and phosphorylates the kidney slit diaphragm component Nephrin. *J. Biol. Chem.* **278**, 20716–20723 (2003).
- Li, H., Lemay, S., Aoudjit, L., Kawachi, H. & Takano, T. SRC-family kinase Fyn phosphorylates the cytoplasmic domain of nephrin and modulates its interaction with podocin. *J. Am. Soc. Nephrol.* **15**, 3006–3015 (2004).
- Lahdenpera, J. *et al.* Clustering-induced tyrosine phosphorylation of nephrin by Src family kinases. *Kidney Int.* **64**, 404–413 (2003).
- Yu, C. C., Yen, T. S., Lowell, C. A. & DeFranco, A. L. Lupus-like kidney disease in mice deficient in the Src family tyrosine kinases Lyn and Fyn. *Curr. Biol.* **11**, 34–38 (2001).
- Putala, H., Sainio, K., Sariola, H. & Tryggvason, K. Primary structure of mouse and rat nephrin cDNA and structure and expression of the mouse gene. *J. Am. Soc. Nephrol.* **11**, 991–1001 (2000).
- Huber, T. B. *et al.* Nephrin and CD2AP associate with phosphoinositide 3-OH kinase and stimulate AKT-dependent signaling. *Mol. Cell. Biol.* **23**, 4917–4928 (2003).
- Chen, M. *et al.* Identification of Nck family genes, chromosomal localization, expression, and signaling specificity. *J. Biol. Chem.* **273**, 25171–25178 (1998).
- Gruenheid, S. *et al.* Enteropathogenic *E. coli* Tir binds Nck to initiate actin pedestal formation in host cells. *Nature Cell Biol.* **3**, 856–859 (2001).
- Lu, W., Katz, S., Gupta, R. & Mayer, B. J. Activation of Pak by membrane localization mediated by an SH3 domain from the adaptor protein Nck. *Curr. Biol.* **7**, 85–94 (1997).
- Rohatgi, R., Nollau, P., Ho, H. Y., Kirschner, M. W. & Mayer, B. J. Nck and phosphatidylinositol 4,5-bisphosphate synergistically activate actin polymerization through the N-WASP-Arp2/3 pathway. *J. Biol. Chem.* **276**, 26448–26452 (2001).
- Bladt, F. *et al.* The murine Nck SH2/SH3 adaptors are important for the development of mesoderm-derived embryonic structures and for regulating the cellular actin network. *Mol. Cell. Biol.* **23**, 4586–4597 (2003).
- Rivera, G. M., Briceno, C. A., Takeshima, F., Snapper, S. B. & Mayer, B. J. Inducible clustering of membrane-targeted SH3 domains of the adaptor protein Nck triggers localized actin polymerization. *Curr. Biol.* **14**, 11–22 (2004).
- Putala, H., Soininen, R., Kilpelainen, P., Wartiovaara, J. & Tryggvason, K. The murine nephrin gene is specifically expressed in kidney, brain and pancreas: inactivation of the gene leads to massive proteinuria and neonatal death. *Hum. Mol. Genet.* **10**, 1–8 (2001).
- Rantanen, M. *et al.* Nephrin TRAP mice lack slit diaphragms and show fibrotic glomeruli and cystic tubular lesions. *J. Am. Soc. Nephrol.* **13**, 1586–1594 (2002).
- Donoviel, D. B. *et al.* Proteinuria and perinatal lethality in mice lacking NEPH1, a novel protein with homology to NEPHRIN. *Mol. Cell. Biol.* **21**, 4829–4836 (2001).
- Ciani, L., Patel, A., Allen, N. D. & French-Constant, C. Mice lacking the giant protocadherin mFAT1 exhibit renal slit junction abnormalities and a partially penetrant cyclopia and anophthalmia phenotype. *Mol. Cell Biol.* **23**, 3575–3582 (2003).
- Kreidberg, J. A. *et al.* Alpha 3 beta 1 integrin has a crucial role in kidney and lung organogenesis. *Development* **122**, 3537–3547 (1996).
- Kaplan, J. M. *et al.* Mutations in *ACTN4*, encoding α -actinin-4, cause familial focal segmental glomerulosclerosis. *Nature Genet.* **24**, 251–256 (2000).
- Welsch, T., Endlich, N., Kriz, W. & Endlich, K. CD2AP and p130Cas localize to different F-actin structures in podocytes. *Am. J. Physiol. Renal Physiol.* **281**, F769–F777 (2001).
- Kos, C. H. *et al.* Mice deficient in α -actinin-4 have severe glomerular disease. *J. Clin. Invest.* **111**, 1683–1690 (2003).
- Frischke, F. *et al.* Actin-based motility of vaccinia virus mimics receptor tyrosine kinase signalling. *Nature* **401**, 926–929 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Huber, G. Rivera, B. Mayer, P. Mundel, P. Soriano, M. Kirschner, K. Tryggvason and D. Hedley for providing reagents and technical advice, D. Holmyard for electron microscopy, and N. Woody for technical contributions. N.J. was supported by fellowships from the Canadian Institutes for Health Research (CIHR) and the National Cancer Institute of Canada (NCIC), with funds from the Terry Fox Run. This work was supported by grants from the CIHR (to S.S.-C.L., T.T., S.E.Q. and T.P.), the NCIC (to S.S.-C.L., S.E.Q. and T.P.) and the Kidney Foundation of Canada (to T.T.). S.S.-C.L. is an NCIC Scientist with funds from the Canadian Cancer Society. L.L. and T.T. hold scholarships from the Fonds de la Recherche en Santé du Québec. S.E.Q. holds a Tier 2 Canada Research Chair. T.P. is a Distinguished Investigator of the CIHR.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.P. (pawson@mshri.on.ca).

LETTERS

Semi-conservative DNA replication through telomeres requires Taz1

Kyle M. Miller^{1*†}, Ofer Rog^{1*} & Julia Promisel Cooper¹

Telomere replication is achieved through the combined action of the conventional DNA replication machinery and the reverse transcriptase, telomerase. Telomere-binding proteins have crucial roles in controlling telomerase activity; however, little is known about their role in controlling semi-conservative replication, which synthesizes the bulk of telomeric DNA¹. Telomere repeats in the fission yeast *Schizosaccharomyces pombe* are bound by Taz1, a regulator of diverse telomere functions^{2–4}. It is generally assumed that telomere-binding proteins impede replication fork progression. Here we show that, on the contrary, Taz1 is crucial for efficient replication fork progression through the telomere. Using two-dimensional gel electrophoresis⁵, we find that loss of Taz1 leads to stalled replication forks at telomeres and internally placed telomere sequences, regardless of whether the telomeric G-rich strand is replicated by leading- or lagging-strand synthesis. In contrast, the Taz1-interacting protein Rap1 is dispensable for efficient telomeric fork progression. Upon loss of telomerase, *taz1Δ* telomeres are lost precipitously, suggesting that maintenance of *taz1Δ* telomere repeats cannot be sustained through semi-conservative replication. As the human telomere proteins TRF1 and TRF2 are Taz1 orthologues, we predict that one or both of the human TRFs may orchestrate fork passage through human telomeres. Stalled forks at dysfunctional human telomeres are likely to accelerate the genomic instability that drives tumorigenesis.

Loss of Taz1 results in dramatic telomere elongation due to deregulation of telomerase^{2,6}. It also causes lethal chromosome-end fusions, but only when the level of nonhomologous end-joining is increased through G1 arrest, a situation not generally experienced by logarithmically growing fission yeast, as their cell cycle lacks a discernible G1 phase³. This unique property of fission yeast allows *taz1Δ* cells to grow despite telomere dysfunction, allowing us to address the involvement of telomere-binding proteins in semi-conservative telomere replication.

To analyse replication intermediates, we subjected genomic DNA to two-dimensional (2D) gel electrophoresis⁵ (described in Supplementary Information and Fig. 1a). *NsiI* digestion released telomere-containing restriction fragments from five of the six wild-type chromosome ends in *S. pombe*, as seen by Southern blotting (Fig. 1b, c). 2D gel electrophoresis revealed that all of these telomere-containing fragments are replicated as simple Y-arcs (Fig. 1d), indicating that replication proceeds unidirectionally from a centromere-proximal origin. This result is consistent with previous work showing that fission yeast replication origins are all markedly (A+T)-rich, in contrast to the (G+C)-rich telomere repeats, and that regions 7–17 kb centromeric to the terminal *NsiI* sites contain multiple replication origins⁷. Notably, some of the wild-type telomere-containing fragments showed replication pause sites

in the subtelomeric region (Fig. 1d), as observed in budding yeast (see below).

Southern blot analysis of *taz1Δ* telomere-containing *NsiI*-digested fragments revealed a heterogeneous smear of products, with most telomere probe hybridization between 5 kb and 7 kb (Fig. 1c). We therefore expected that 2D gel electrophoresis of DNA from *taz1Δ* cells would yield an amalgamation of superimposed Y-arcs ranging from ≤5 kb for the smallest 1N spots (unreplicated DNA) to ≥14 kb for the largest 2N spots (almost completely replicated DNA), with the vast majority of signal appearing as a semicircle encompassed by arcs F3 and F5 (Fig. 1d). Instead, the telomere signal intensified at the top of the superimposed Y-arcs and spread into higher molecular weight fragments, forming a plume that extended well beyond the predicted size for the replicating telomeres (Fig. 1d). This plume of intermediates of higher molecular weight and complexity was not accompanied by corresponding intermediates descending towards the arc of linears. This anomalous migration pattern is due neither to an inability of our gel system to detect Y-arcs representing long DNA fragments, nor to an artefact generated by *in vitro* interactions between the extensive G-rich overhangs at *taz1Δ* telomeres (Supplementary Fig. 1). Thus, the lack of appreciable signal descending from the top of the arcs to the 2N region suggests that forks rarely progress to the ends of *taz1Δ* telomeres.

Rap1 binds to Taz1 and mediates a subset of its functions^{8–10}. Hence, telomeres of long and heterogeneous length, as well as increased telomeric 3' overhang signals, are found in both *taz1Δ* and *rap1Δ* cells^{8–10} (Fig. 1e). Strikingly, however, *rap1Δ* telomeres did not share the anomalous replication pattern of *taz1Δ* telomeres (Fig. 1e). Rather, *rap1Δ* telomeres yielded discernable broad Y-arcs, with smears of hybridization descending from the top of the superimposed Y-arcs to the arc of linears, suggesting that replication forks proceed to the terminus of a substantial fraction of *rap1Δ* telomeres. The 2D gel electrophoresis pattern suggests that not all *rap1Δ* telomeres are fully replicated, consistent with the idea that cellular Taz1 levels are insufficient to bind the entire *rap1Δ* telomere, effectively leading to 'partially *taz1Δ*' telomeres. Nevertheless, the clear difference between the *taz1Δ* and *rap1Δ* 2D gel electrophoresis patterns indicates that the *taz1Δ* pattern is not merely a consequence of increased telomere length or length-heterogeneity.

To assess replication of telomeric DNA without the complication of heterogeneous telomere size, we analysed an internally placed telomeric tract of 250 bp, inserted ~8 kb from a well-characterized cluster of replication origins on chromosome III (Fig. 2a; A. Hebden, K.M.M. and J.P.C., manuscript in preparation)¹¹. As expected, *NsiI* digestion yielded a 3.5-kb internal telomere-repeat-containing fragment that was replicated as a simple Y-arc in wild-type cells. However, *taz1⁺* deletion resulted in fork stalling specifically through the region containing the telomere repeat stretch (Fig. 2a and

¹Telomere Biology Laboratory, Cancer Research UK, 44 Lincoln's Inn Fields, London WC2A 3PX, UK. [†]Present address: Department of Biochemistry and Biophysics, Cancer Research Institute, University of California, San Francisco, California 94115-0128, USA.

*These authors contributed equally to this work.

Supplementary Fig. 2). Thus, telomeric DNA lacking Taz1 results in fork pausing regardless of whether the telomere stretch terminates with a chromosome end.

A sizeable fraction of the molecules present in these Y-arcs completes replication, indicating that pausing at these internal telomeric sites is reversible or does not occur in all *taz1Δ* cells. As we did not observe a difference in the growth profiles of *taz1⁺* versus *taz1Δ* strains carrying internal telomere stretches (data not shown), we infer that stalled forks at these internal stretches do not elicit a pronounced cell-cycle arrest, consistent with the observation that pausing at the replication fork barrier RTS1 fails to trigger measurable cell-cycle arrest¹². Although the probability of pausing may be finite over a 250-bp internal telomere stretch, the chances of a fork progressing to the end of a 1–4-kb *taz1Δ* telomere decrease exponentially with distance from the centromere. Accumulation of stalled forks along a *taz1Δ* telomere may underlie the anomalous replication patterns seen at *taz1Δ* chromosome ends (Fig. 1d, e; discussed below).

Natural telomeres are oriented such that the G-rich telomeric strand always serves as the template for lagging-strand synthesis, and the internal telomere sequence shown in Fig. 2a shares this polarity. To determine whether the telomeric replication fork pausing seen in the absence of Taz1 is linked specifically to lagging-strand replication of the G-rich strand, we next oriented the internal telomere stretch in

reverse to natural telomeres, such that the G-rich strand is copied by leading-strand replication (Fig. 2b). Deletion of *taz1⁺* from this strain again resulted in the accumulation of stalled replication forks in the telomeric stretch. Therefore, Taz1 is required for efficient replication of telomeric repeats regardless of their orientation with respect to the direction of replication.

To explore the extent of anomalous replication of *taz1Δ* telomeres, we examined replication in subtelomeric regions. *ApaI* sites, located as little as 32 bp from telomeres¹³, can be used to separate subtelomeric sequences from telomeres (Fig. 1b). Analysis of the subtelomeric region revealed four *ApaI* fragments (Fig. 3a), two of which could be analysed by 2D gel electrophoresis. As expected, subtelomeric fragments are replicated as simple Y-arcs (Fig. 3b). The left (descending) portion of the Y-arc comprises replication intermediates nearing completion (hence the ~2N DNA content), at which point each intermediate is resolved to two 1N DNA molecules. However, *taz1Δ* cells accumulated paused replication forks at the position represented by the 2N spot on arc F2 (arrow, Fig. 3b), which lies adjacent to the telomeric tracts (Fig. 1b). Although additional spots can be seen along the arc of linears, they vary between experiments whereas the 2N spots are seen reproducibly. Therefore, Taz1 is required for efficient fork progression specifically through the distal, telomeric end of the subtelomeric region. As unwinding activities associated with the replication fork may occur up to

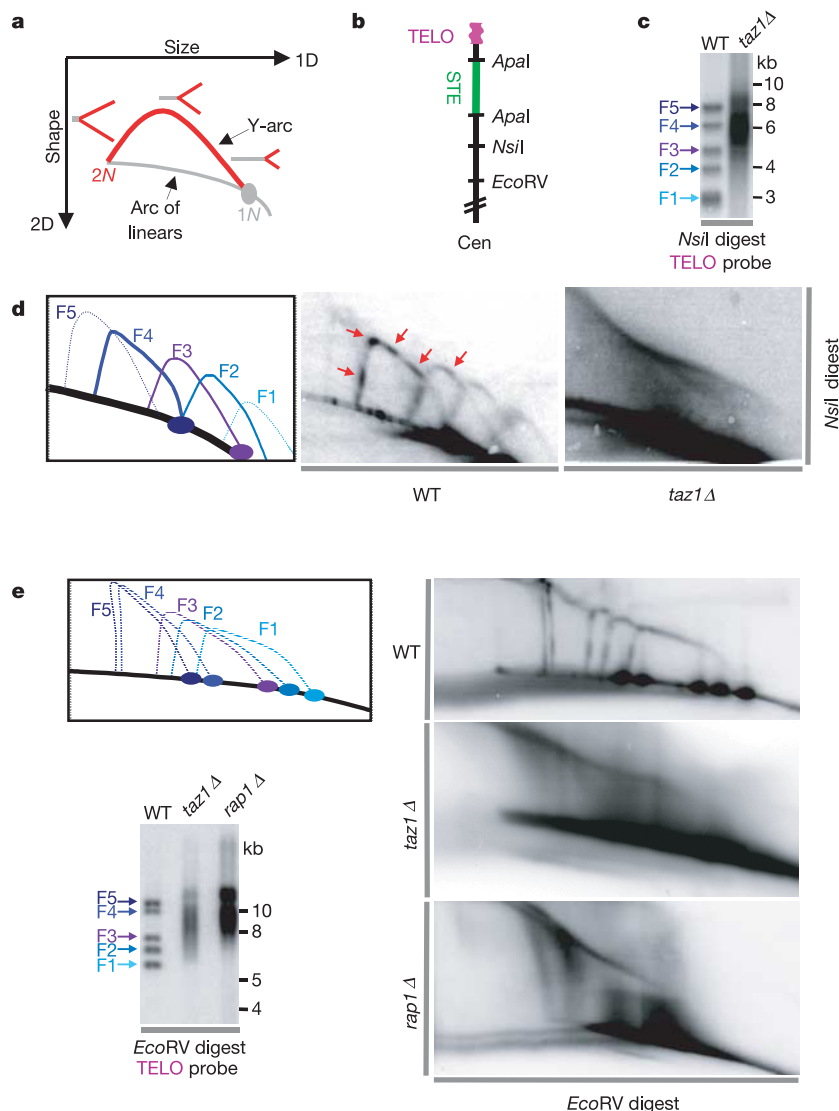


Figure 1 | Efficient semi-conservative replication through telomeres requires Taz1 but not Rap1.

a, The Y-arc pattern seen by 2D gel electrophoresis. For each restriction fragment, the unreplicated form runs as the '1N spot'. The '2N spot' represents the theoretical duplication of that fragment. Unreplicated DNA fragments of all sizes run in the arc of linears. The Y-arc pattern is generated by unidirectional movement of a replication fork across a DNA fragment. Unreplicated DNA is shown in grey, replicated DNA in red. **b**, Relative positions of restriction sites in telomeric (TELO) and subtelomeric (STE) regions of chromosomes I and II. Cen, centromere. **c**, Southern blot analysis of telomeric *NsiI*-digested fragments. **d**, 2D gel analysis of telomeric *NsiI*-digested fragments. 'F' labels on the interpretive diagram indicate the corresponding unreplicated bands in **c**. Red arrows indicate subtelomeric pause sites. **e**, *taz1Δ* and *rap1Δ* cells show distinct patterns of telomere replication. Southern blot (left) and 2D gel (right) of telomeric *EcoRV*-digested fragments.

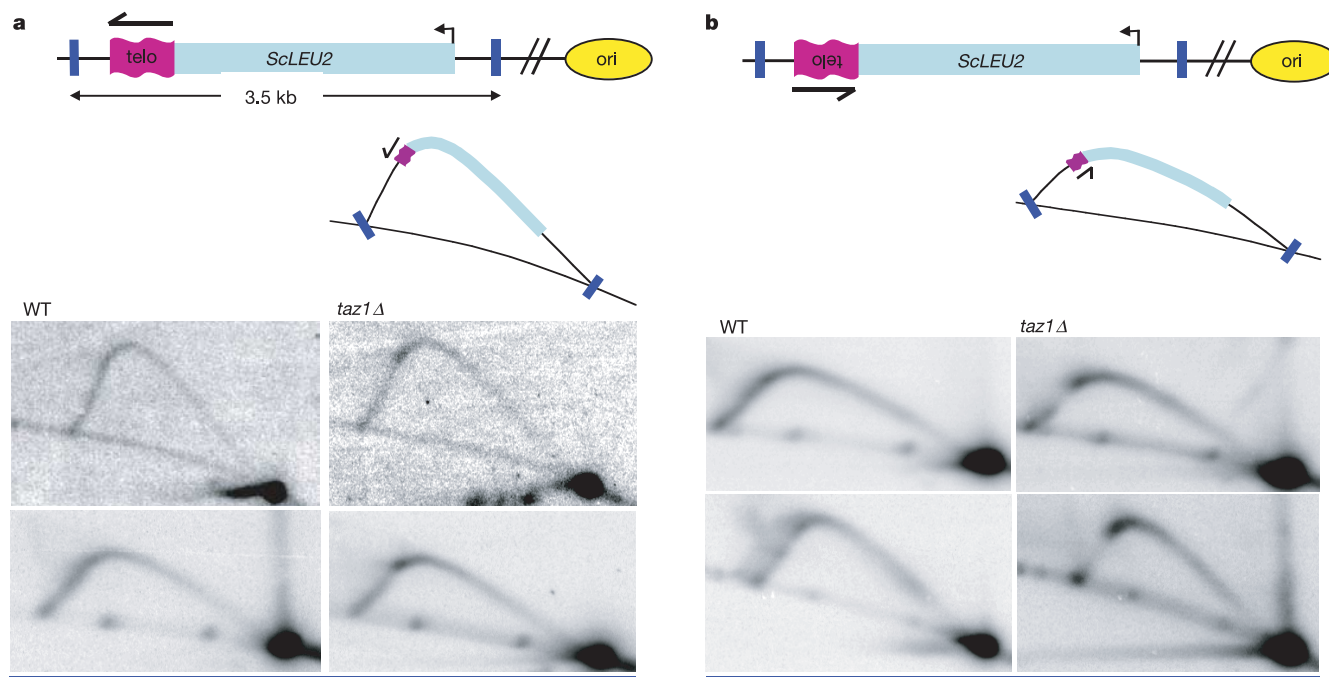


Figure 2 | Taz1 is required for efficient replication through internally placed telomere tracts. **a, b,** Maps showing the genomic region of *S. pombe* chromosome III containing an ectopically inserted telomere tract. The ARS3002-4 cluster (ori) is located ~8 kb centromeric to *ura4* and has been shown to replicate this locus¹¹. In **b**, synthetic telomere constructs of 200 bp (top blots) and 450 bp (bottom blots) were inserted at the *ura4* locus as in **a**,

but in the opposite orientation, with the 3' end of the telomeric G-rich strand oriented towards ori. The telomere tract is 2.5 kb from the ori-proximal end of the 3.5-kb *NsiI* fragment. In each illustration, the telomeric stretch is shown in purple, and the position of *ScLEU2* is shown in light blue. *NsiI*-digested DNA was subjected to 2D gel electrophoresis and hybridized with a *ScLEU2* probe.

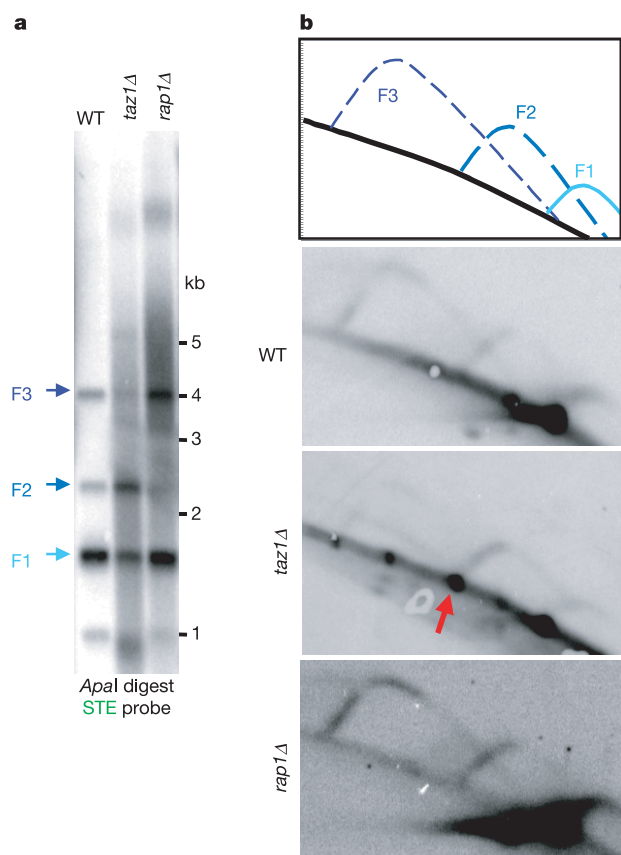


Figure 3 | Taz1 is required for replication through the telomere/subtelomere boundary. **a,** Southern blot probed for subtelomeric sequences (STE). **b,** 2D gel electrophoresis of *Apal*-digested DNA, probed for STE. The 1N spot of F3 overlaps the end of F1, obscuring its analysis. Red arrow indicates the 2N spot on F2.

~200 bp ahead of the fork itself^{14,15}, we speculate that in the absence of Taz1, these activities are obstructed as they encounter the telomere, leading to stalling at the telomere/subtelomere boundary. Notably, *rap1⁺* deletion did not elicit replication fork pausing at the distal end of the subtelomeric region (Fig. 3b). Thus, Taz1 binding to telomeres is itself sufficient to prevent fork stalling at the subtelomere/telomere boundary, and this protection does not require Rap1.

If conventional replication forks are unable to efficiently traverse *taz1Δ* telomeres, then most of the 1–4-kb *taz1Δ* telomeres must be synthesized by telomerase in every cell cycle. If this were true, deletion of the gene encoding the telomerase reverse transcriptase (*trt1⁺*) in a *taz1Δ* background would trigger a precipitous loss of telomere repeats, rather than the gradual telomere attrition seen upon *trt1⁺* deletion in wild-type cells. Previous results support this prediction, as loss of Taz1 in cells lacking telomerase has been shown to accelerate viability loss by 100-fold^{6,16}—an initially puzzling observation given that the elongated *taz1Δ* telomeres might be expected to postpone senescence. To address this further, we monitored the fates of long telomeres after *trt1⁺* deletion in *taz1Δ* and *rap1Δ* backgrounds by sporulating *taz1Δ trt1^{+/Δ}* and *rap1Δ trt1^{+/Δ}* diploids and following telomere length over time. Both parental diploid strains had highly elongated telomeres, as expected (Fig. 4a, b). However, a marked difference between *taz1Δ* and *rap1Δ* telomeres was seen in spores lacking Trt1. When single colonies from *rap1Δ trt1Δ* spores were continuously cultured, their telomeres showed progressive length reduction, culminating in near-complete loss of telomere signal by ~150 generations (Fig. 4a). In contrast, *taz1Δ trt1Δ* colonies showed erratic telomere hybridization patterns, with an immediate loss of the bulk of the telomere signal followed by a variable level of signal maintenance. Similar phenomena were observed when spores were germinated *en masse* (under selection for the *trt1⁺* deletion) to allow for the extraction of sufficient quantities of DNA for Southern blot analysis at earlier time points (starting at ~10 generations)—telomere length declined progressively in *rap1Δ trt1Δ* populations but showed an abrupt decline

followed by variable maintenance in *taz1Δ trt1Δ* populations (Fig. 4b). The abrupt telomere loss seen upon deletion of *trt1*⁺ from *taz1Δ* cells is consistent with an inability to maintain *taz1Δ* telomeres through conventional replication. Likewise, the gradual telomere attrition seen in *rap1Δ trt1Δ* cells agrees with our observation that semi-conservative telomere replication is largely complete in the absence of Rap1.

Replication fork pausing can trigger not only DNA breakage, but also DNA recombination pathways^{12,17,18}. The distinct patterns of telomere attrition and survival of telomerase loss in *taz1Δ* and *rap1Δ* backgrounds are compatible with the idea that high levels of telomeric fork stalling are unique to *taz1Δ* cells and promote telomeric recombination. *rap1Δ trt1Δ* strains, like 'normal' *trt1Δ* strains, eventually lose all telomeric DNA and survive by sustaining intramolecular end-fusions that result in circularization of each of the three chromosomes (Fig. 4c). In contrast, *taz1Δ trt1Δ* survivors maintain variable amounts of telomeric DNA, presumably through recombination. The uniqueness of this survival mode to *taz1Δ*, and not *rap1Δ*, cells supports the idea that it stems from stalled forks that elicit hyper-recombination.

Paused replication at *taz1Δ* telomeres may generate unwound, unreplicated strands of DNA that can invade other chromosomes. The G-rich nature of telomere repeats and/or long 3' single-strand overhangs in *taz1Δ* cells might facilitate such interactions between strands of DNA from distinct duplexes. Thus, the plumes

of hybridization seen upon 2D gel electrophoresis of *taz1Δ* telomeres may represent branched structures generated from telomeric strand invasion reactions. Intriguingly, *taz1Δ* mutants accumulate entangled telomeres that fail to separate at mitosis if the preceding S-phase occurred at a cold temperature⁴ ($\leq 20^\circ\text{C}$). Thus, we speculate that stalled replication forks at *taz1Δ* telomeres elicit telomeric entanglement, manifest by the plume pattern in 2D gels, and that resolution of these entangled telomeres is compromised in the cold. Like the precipitous telomere loss upon *trt1*⁺ deletion, the *taz1Δ* chromosomal entanglement phenotype is absent from *rap1Δ* cells¹⁰, supporting the idea that stalled telomeric replication forks are the originating defect.

How might Taz1 facilitate replication through telomere sequences? Taz1 may coordinate loosening, remodelling or removal of the telomere complex to allow replication fork progression. Alternatively, Taz1 may recruit factors that facilitate unwinding of the G-rich telomere repeats or prevent them from forming spurious intra- or interstrand G-rich quadruplex structures once unwound. Such challenges have been invoked to explain defective telomere replication in mammalian cells lacking the WRN or RTEL helicases^{19,20}. The latter replication defects were attributed to exposure of G-rich single-stranded DNA during lagging-strand telomere replication, reflecting the disposition of natural telomeres. By creating an artificial situation in which the telomeric G-rich strand is replicated by leading-strand synthesis, we found that both leading- and lagging-strand replication

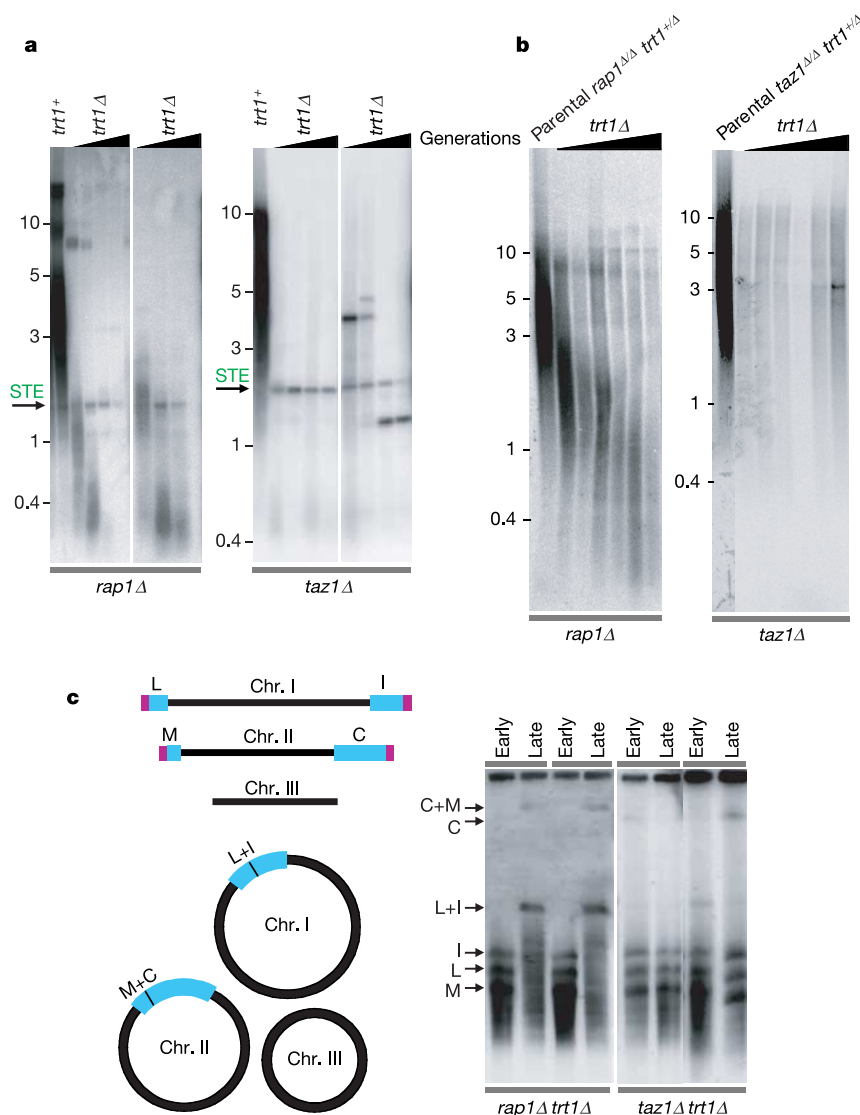


Figure 4 | Differences in the kinetics of telomere attrition and survival mechanism in *taz1Δ* versus *rap1Δ* cells after *trt1*⁺ deletion. **a**, Diploids homozygous for *taz1Δ* or *rap1Δ* and heterozygous for *trt1Δ* were sporulated. *trt1Δ* progeny were selected, serially re-streaked and subjected to Southern blot analysis of telomere length every 25–30 generations. **b**, The diploids described in **a** were sporulated and germinated *en masse*. Cultures were grown logarithmically under selection for *trt1Δ* and analysed every 24 h (~5–10 generations). **c**, DNA from 'early' (~50 generations after *trt1* loss) and 'late' (~200 generations after *trt1* loss) cultures of *trt1Δ* survivors was digested with *NotI*, separated by pulsed-field gel electrophoresis, blotted and hybridized with probes against 'L', 'M', 'I' and 'C' (left panel) on chromosomes (Chr.) I and II. L, M, I and C bands indicate linear chromosomes; L+I and M+C fusion bands indicate circular chromosomes.

of any parental telomeric G-rich strand requires Taz1; this may also apply to the requirement for WRN and RTEL. These studies raise the possibility that Taz1, and by analogy human TRF1 or TRF2, may promote unimpeded replication by recruiting such helicases to telomeres.

Previous studies have suggested that telomere-binding proteins hinder, rather than facilitate, replication fork progression. Stalled forks have been observed at sites of steric hindrance by stable DNA–protein complexes, including ribosomal DNA and telomeric regions in budding yeast^{21,22}. Deletion of the Rap1 carboxy terminus had no effect on replication pausing through budding yeast telomeres, while shorter telomeres alleviated pausing²³. These results led to the conclusion that the Rap1 amino terminus, which contains its DNA-binding domain, generates an obstacle to replication fork passage, with the larger number of bound Rap1 molecules at long telomeres exacerbating this effect²³. However, loss of the N-terminal DNA-binding domain of Rap1 would be more equivalent to *taz1*⁺ deletion, and might reveal a role for Rap1 in promoting, rather than impeding, fork passage. In an *in vitro* simian virus 40 (SV40) replication system, expression of human TRF1 and TRF2 creates a replication block at the telomeric termini of a linear plasmid²⁴. We suggest that, by analogy with Taz1, physiological concentrations of human TRFs may facilitate, rather than hinder, passage of the replication fork *in vivo*.

Our results suggest that stalled forks may arise from an altered balance of telomere-binding proteins accompanying telomere attrition in human telomerase-negative cells. Stalled replication forks may accelerate telomere attrition and/or trigger the recombination reactions that constitute the ‘alternative lengthening of telomeres (ALT)’^{25,26} and ‘telomere rapid deletion (TRD)’^{27,28} pathways seen in telomerase-negative cancer cells. Hence, the ‘end replication problem’ begins before the replication fork reaches the chromosomal terminus, and necessitates telomere-binding proteins not only to recruit and control telomerase but also to ensure passage of the semi-conservative replication fork.

METHODS

All experiments were performed using *Schizosaccharomyces pombe* strains described in the Supplementary Information. Replication intermediates were detected using two-dimensional neutral–neutral agarose gel electrophoresis. The positions of various replication intermediates are indicated in each figure. To test telomere length after loss of Trt1, diploid strains were sporulated and the DNA extracted from the spores and their progeny was subjected to Southern blotting and hybridization with telomere probes. For details of the methods used, see Supplementary Information.

Received 13 August 2005; accepted 9 February 2006.

- Teixeira, M. T., Arneric, M., Sperisen, P. & Lingner, J. Telomere length homeostasis is achieved via a switch between telomerase-extendable and -nonextendable states. *Cell* **117**, 323–335 (2004).
- Cooper, J. P., Nimmo, E. R., Allshire, R. C. & Cech, T. R. Regulation of telomere length and function by a Myb-domain protein in fission yeast. *Nature* **385**, 744–747 (1997).
- Ferreira, M. G. & Cooper, J. P. The fission yeast Taz1 protein protects chromosomes from Ku-dependent end-to-end fusions. *Mol. Cell* **7**, 55–63 (2001).
- Miller, K. M. & Cooper, J. P. The telomere protein Taz1 is required to prevent and repair genomic DNA breaks. *Mol. Cell* **11**, 303–313 (2003).
- Brewer, B. J. & Fangman, W. L. The localization of replication origins on ARS plasmids in *S. cerevisiae*. *Cell* **51**, 463–471 (1987).
- Nakamura, T. M., Cooper, J. P. & Cech, T. R. Two modes of survival of fission yeast without telomerase. *Science* **282**, 493–496 (1998).
- Segurado, M., de Luis, A. & Antequera, F. Genome-wide distribution of DNA replication origins at A+T-rich islands in *Schizosaccharomyces pombe*. *EMBO Rep.* **4**, 1048–1053 (2003).
- Kanoh, J. & Ishikawa, F. spRap1 and spRif1, recruited to telomeres by Taz1, are essential for telomere function in fission yeast. *Curr. Biol.* **11**, 1624–1630 (2001).
- Chikashige, Y. & Hiraoka, Y. Telomere binding of the Rap1 protein is required for meiosis in fission yeast. *Curr. Biol.* **11**, 1618–1623 (2001).
- Miller, K. M., Ferreira, M. G. & Cooper, J. P. Taz1, Rap1 and Rif1 act both interdependently and independently to maintain telomeres. *EMBO J.* **24**, 3128–3185 (2005).
- Dubey, D. D., Zhu, J., Carlson, D. L., Sharma, K. & Huberman, J. A. Three ARS elements contribute to the *ura4* replication origin region in the fission yeast, *Schizosaccharomyces pombe*. *EMBO J.* **13**, 3638–3647 (1994).
- Lambert, S., Watson, A., Sheedy, D. M., Martin, B. & Carr, A. M. Gross chromosomal rearrangements and elevated recombination at an inducible site-specific replication fork barrier. *Cell* **121**, 689–702 (2005).
- Sugawara, N. *DNA Sequences at the Telomeres of the Fission Yeast S. pombe*. PhD thesis, Harvard Univ. (1988).
- Sogo, J. M., Lopes, M. & Foiani, M. Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects. *Science* **297**, 599–602 (2002).
- Gasser, R., Koller, T. & Sogo, J. M. The stability of nucleosomes at the replication fork. *J. Mol. Biol.* **258**, 224–239 (1996).
- Beernink, H. T., Miller, K., Deshpande, A., Bucher, P. & Cooper, J. P. Telomere maintenance in fission yeast requires an *Est1* ortholog. *Curr. Biol.* **13**, 575–580 (2003).
- Cotta-Ramusino, C. *et al.* Exo1 processes stalled replication forks and counteracts fork reversal in checkpoint-defective cells. *Mol. Cell* **17**, 153–159 (2005).
- Ahn, J. S., Osman, F. & Whitby, M. C. Replication fork blockage by *RTS1* at an ectopic site promotes recombination in fission yeast. *EMBO J.* **24**, 2011–2023 (2005).
- Crabbe, L., Verdun, R. E., Haggblom, C. I. & Karlseder, J. Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science* **306**, 1951–1953 (2004).
- Ding, H. *et al.* Regulation of murine telomere length by Rtel: an essential gene encoding a helicase-like protein. *Cell* **117**, 873–886 (2004).
- Ivessa, A. S. *et al.* The *Saccharomyces cerevisiae* helicase Rrm3p facilitates replication past nonhistone protein–DNA complexes. *Mol. Cell* **12**, 1525–1536 (2003).
- Ivessa, A. S., Zhou, J. Q., Schulz, V. P., Monson, E. K. & Zakian, V. A. *Saccharomyces Rrm3p*, a 5′ to 3′ DNA helicase that promotes replication fork progression through telomeric and subtelomeric DNA. *Genes Dev.* **16**, 1383–1396 (2002).
- Makovets, S., Herskowitz, I. & Blackburn, E. H. Anatomy and dynamics of DNA replication fork movement in yeast telomeric regions. *Mol. Cell. Biol.* **24**, 4019–4031 (2004).
- Ohki, R. & Ishikawa, F. Telomere-bound TRF1 and TRF2 stall the replication fork at telomeric repeats. *Nucleic Acids Res.* **32**, 1627–1637 (2004).
- Londono-Vallejo, J. A., Der-Sarkissian, H., Cazes, L., Bacchetti, S. & Reddel, R. R. Alternative lengthening of telomeres is characterized by high rates of telomeric exchange. *Cancer Res.* **64**, 2324–2327 (2004).
- Laud, P. R. *et al.* Elevated telomere–telomere recombination in WRN-deficient, telomere dysfunctional cells promotes escape from senescence and engagement of the ALT pathway. *Genes Dev.* **19**, 2560–2570 (2005).
- Lustig, A. J. Clues to catastrophic telomere loss in mammals from yeast telomere rapid deletion. *Nature Rev. Genet.* **4**, 916–923 (2003).
- Wang, R. C., Smogorzewska, A. & de Lange, T. Homologous recombination generates T-loop-sized deletions at human telomeres. *Cell* **119**, 355–368 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank our laboratory members for discussions and support, and are particularly grateful to A. Hebden for sharing the internal telomere strain, J. Lingner, S. Marcand, Y. Mazon and F. Uhlmann for discussions, and B. Arcangioli, A. Kaykov and C. Heichinger for advice on 2D gel electrophoresis. This work was supported by Cancer Research UK.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.P.C. (julie.cooper@cancer.org.uk).

Proton-coupled electron transfer drives the proton pump of cytochrome *c* oxidase

Ilya Belevich¹, Michael I. Verkhovsky¹ & Mårten Wikström¹

Electron transfer in cell respiration is coupled to proton translocation across mitochondrial and bacterial membranes, which is a primary event of biological energy transduction. The resulting electrochemical proton gradient is used to power energy-requiring reactions, such as ATP synthesis. Cytochrome *c* oxidase is a key component of the respiratory chain, which harnesses dioxygen as a sink for electrons and links O₂ reduction to proton pumping¹. Electrons from cytochrome *c* are transferred sequentially to the O₂ reduction site of cytochrome *c* oxidase via two other metal centres, Cu_A and haem *a*, and this is coupled to vectorial proton transfer across the membrane by a hitherto unknown mechanism. On the basis of the kinetics of proton uptake and release on the two aqueous sides of the membrane, it was recently suggested that proton pumping by cytochrome *c* oxidase is not mechanistically coupled to internal electron transfer². Here we have monitored translocation of electrical charge equivalents as well as electron transfer within cytochrome *c* oxidase in real time. The results show that electron transfer from haem *a* to the O₂ reduction site initiates the proton pump mechanism by being kinetically linked to an internal vectorial proton transfer. This reaction drives the proton pump and occurs before relaxation steps in which protons are taken up from the aqueous space on one side of the membrane and released on the other².

Reduction of O₂ requires four electrons. One electron is transferred at a time to the binuclear haem *a*₃/Cu_B site of O₂ reduction, which therefore attains several intermediate states during the catalytic cycle. These states may be distinguished kinetically and their structures assigned by spectroscopic methods^{3–5}. O₂ reduction by cytochrome *c* oxidase is coupled to pumping of one proton across the membrane for each of the four transferred electrons¹. In addition, one 'substrate proton' per electron is taken up into the binuclear site from the negatively charged N-side of the membrane to form the product water. In the early stages of the catalytic cycle studied here, uptake of both substrate and pumped protons takes place via the D-pathway⁵, which is named after the residue D124, and ends at the residue E278 in the middle of the membrane (Fig. 1a; amino acid numbering corresponds to subunit I of cytochrome *c* oxidase from *Paracoccus denitrificans*).

Some structural features of cytochrome *c* oxidase are similar to those of the light-driven proton pump bacteriorhodopsin⁶. A recent study proposed a much closer similarity by concluding that the mechanism of proton pumping in cytochrome *c* oxidase is not kinetically coupled to electron transfer, but rather occurs as a result of conformational changes set forth by uptake of the substrate proton into the binuclear site². This issue is indeed best explored by studying the earliest part of the catalytic cycle² that immediately follows the reaction of the fully reduced enzyme with O₂ (Fig. 1b), because these reactions are unique in that electron transfer from haem *a* into the binuclear site (during A → P_R) is kinetically distinguishable from the next step (P_R → F)^{3–5,7} during which there is both net proton uptake

and pumping^{2,8,9}. These electron and proton transfers are kinetically indistinguishable in all later stages of the cycle. The kinetic separation may thus provide a fulcrum for a deeper general understanding of the proton pump mechanism, because it is reasonable to assume that the

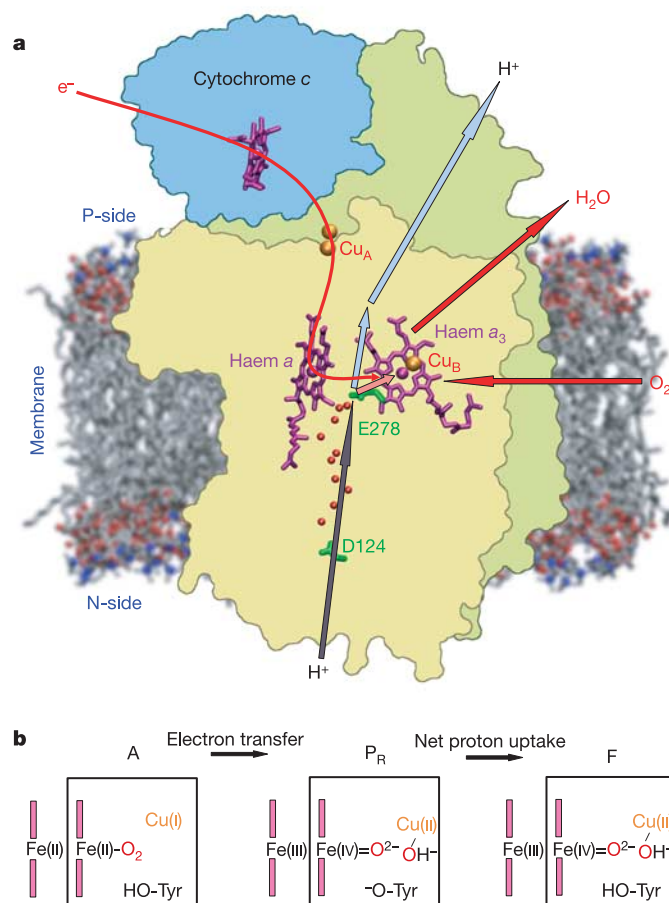


Figure 1 | Structure and function of cytochrome *c* oxidase. **a**, Subunits I (yellow) and II (green) are depicted in the phospholipid membrane, together with docked cytochrome *c* (blue, ref. 25). Proton transfer from the N-side of the membrane via the D-pathway (grey arrow) leads to E278. E278 donates protons, both to the haem *a*₃/Cu_B site to produce water from reduced O₂ (light red arrow), and for pumping across the membrane (blue arrows). The red spheres are crystallographically observed water molecules in the D-pathway (Protein Data Bank number 1v54, ref. 26). The VMD program²⁷ was used in producing the figure. **b**, Early intermediates in the reaction of fully reduced cytochrome *c* oxidase with O₂. The O₂ reduction site (square) includes haem *a*₃, Cu_B and a conserved tyrosine (Tyr)^{17,28,29}. Haem *a* is shown on the left of the square.

¹Helsinki Bioenergetics Group, Institute of Biotechnology, University of Helsinki, FIN-00014 University of Helsinki, Helsinki, Finland.

same basic mechanism is repeated four times during the catalytic cycle.

In the early reaction sequence (Fig. 1b), the reduced binuclear site first binds an O_2 molecule ($\tau \approx 10 \mu s$) to form the oxygen adduct, compound A^{3-5} . Then, there is electron transfer from haem a into the binuclear site with scission of the O–O bond ($\tau \approx 30 \mu s$), leading to a ferryl/cupric intermediate that has been called P_R^7 . P_R is unstable and converts to the F state in the next step ($\tau \approx 80 \mu s$), which is linked to net uptake of a substrate proton from the N-side but involves no further electron transfer into the binuclear site^{2,5,10}. The earliest proton pumping event after O_2 binding has so far been ascribed to this latter reaction^{2,8,9} (but see below). The catalytic cycle continues from F to the oxidized O state ($\tau \approx 1-3$ ms) with transfer of a second electron into the binuclear site, pumping of one proton, and net uptake of another. Regeneration of the reduced binuclear site for binding the next O_2 molecule requires input of two more electrons, and these electron transfers are also each associated with net proton uptake¹¹ and pumping^{9,12}.

Elucidation of the mechanism of proton translocation may require monitoring of proton transfer within the enzyme structure in real time to complement measurements of proton uptake and release². Here we apply a sensitive electrometric technique⁸ to phospholipid vesicles inlaid with cytochrome c oxidase to monitor charge translocation within the membrane-bound enzyme. To determine its kinetic linkage to the catalytic chemistry, reaction intermediates are monitored by time-resolved optical spectroscopy. Two major kinetically distinguishable phases of charge translocation are observed at neutral pH (Fig. 2a) that roughly correspond to the spectroscopically observed kinetics of the $P_R \rightarrow F$ (Fig. 1b) and $F \rightarrow O$ transitions⁸. Each phase is mainly due to proton transfer from the aqueous N-side of the membrane into the binuclear site, and to proton pumping, with only minor contributions from electron transfer. Electron transfer from haem a into the binuclear site does not give rise to membrane potential due to the location of these centres at the same depth within the membrane¹³. The reaction $A \rightarrow P_R$ (Fig. 1b) has indeed been thought to be electrically silent⁸. However, the large amplitude of charge translocation during the subsequent $P_R \rightarrow F$ reaction might overwhelm possible charge translocation during $A \rightarrow P_R$. To explore this, the reaction was studied at high pH, where the $P_R \rightarrow F$ and $F \rightarrow O$ steps are known to be decelerated¹⁴. As shown in Fig. 2a, charge translocation during $F \rightarrow O$ is indeed slowed down. Paradoxically, however, the rate of the fast electrometric phase appears to be accelerated and the initial lag phase shortened (Fig. 2b). In fact, at high pH the fast phase coincides with the kinetics of the prior reaction step, $A \rightarrow P_R$, as monitored spectroscopically (Fig. 2c). Such an apparent acceleration can be explained if the fast phase of potential generation consists of two electrogenic events, the amplitudes of which are differently affected by pH: the amplitude during $A \rightarrow P_R$ is pH-independent, but during $P_R \rightarrow F$ the initially large amplitude decreases with pH. The transient during $A \rightarrow P_R$ is therefore very difficult to discern at neutral pH. Kinetic simulations indeed show that charge translocation during $A \rightarrow P_R$ is compatible with the data also at neutral pH (Fig. 2a, b, dashed lines), but independent evidence is obviously required. We therefore tested the D124N mutant enzyme where proton transfer to E278 via the D-pathway is strongly inhibited (Fig. 1a, ref. 15), and where proton pumping is abolished¹⁶. Therefore, the amplitude of the major electrometric phase is decreased, and the early transient during $A \rightarrow P_R$ should appear if it is present at neutral pH. As shown in Fig. 2b, charge translocation in this mutant enzyme is superimposable on the high pH trace of wild-type enzyme. Hence, an early phase of charge translocation indeed takes place during the $A \rightarrow P_R$ reaction also at neutral pH. Whether the electron transfer during $A \rightarrow P_R$ is essential can be assessed by studying the O_2 reaction of the 'mixed valence' enzyme, where only the binuclear site is initially reduced and no electron transfer from haem a is possible. Here, scission of the O–O bond occurs as before, but by reactions entirely

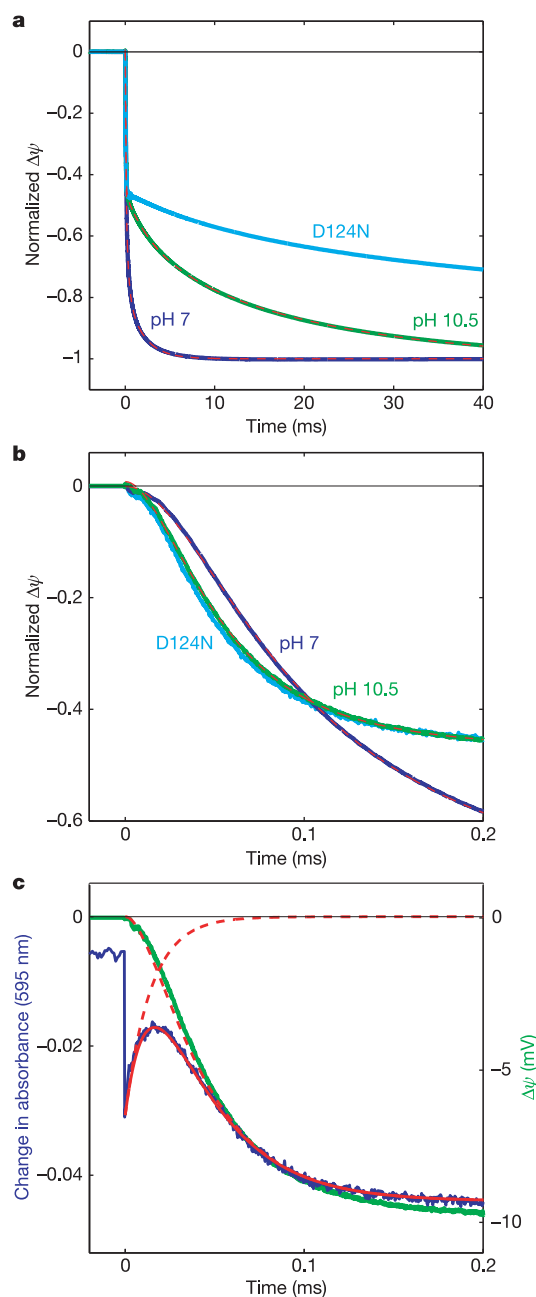


Figure 2 | Charge translocation and electron transfer during the reaction of fully reduced cytochrome c oxidase with O_2 . **a**, Normalized electrometric response of the reconstituted wild-type enzyme at pH 7 (blue), pH 10.5 (green) and with the D124N mutant enzyme at pH 7 (cyan). Time zero corresponds to the moment of the laser flash (see Methods). Red dashed lines on the pH 7 and 10.5 traces represent the theoretical fit of the data with a model of five sequential reaction steps⁸, where the rate constants for the formation and decay of compound A (Fig. 1b) were fixed at the values obtained from the optical data in **c**. **b**, The same traces as in **a** expanding the kinetics of the fast phase. **c**, Comparison of the time courses of formation and decay of compound A (Fig. 1b) at 595 nm (blue trace), with generation of electric potential (green trace), both at pH 10.5. The initial downward deflection of the blue trace is due to CO photo-dissociation. The solid red trace is a two-component exponential fit to the optical data at 595 nm; the dashed red lines show the two components, of which the second (decay of compound A) coincides with the electrometric trace. $\Delta\psi$, change in membrane potential.

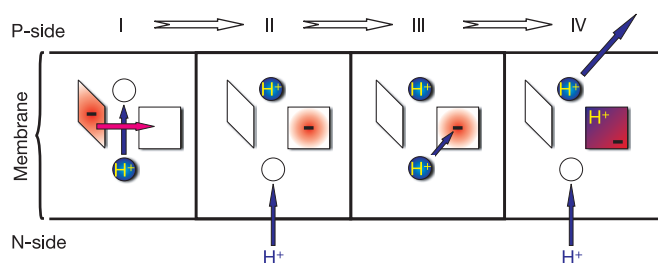


Figure 3 | Scheme of the proposed proton pump mechanism. Four states are shown (I–IV), each comprising haem *a* (rhombus) and the haem *a*₃/Cu_B site (square). The lower and upper circles denote the carboxylic residue E278 at the end of the D-pathway (Fig. 1a) and an unidentified protonatable site above the haems, respectively. In I → II electron transfer from haem *a* to the binuclear site is coupled to transfer of a proton from E278 to the protonatable site. In II → III, E278 is re-protonated from the N-side via the D-pathway. In III → IV, a substrate proton is transferred from E278 to the binuclear site. After IV, E278 is again re-protonated and the proton above the haems is ejected towards the P-side.

within the binuclear site¹⁷. As shown previously⁸, this reaction is not associated with any charge translocation. We therefore conclude that the early charge translocation during A → P_R not only coincides kinetically with electron transfer from haem *a* to the binuclear site, but also depends on it.

At neutral pH the amplitude of charge translocation during formation of the F state corresponds to translocation of about 1.8 charge equivalents across the membrane¹⁸. Vectorial charge movement during A → P_R (approximately 15% of the amplitude at neutral pH) thus amounts to transfer of one electrical charge across about 30% of the membrane dielectric. Because it cannot be due to transmembrane electron transfer, it probably reflects an internal proton transfer reaction towards the positively charged P-side. The residue E278 at the end of the D-pathway is normally protonated before the reaction with O₂. This residue is involved in proton pumping¹⁹ as well as in proton transfer to the binuclear site¹⁵, and is thus a candidate for being the proton donor. This is supported by electrometric experiments with the E278Q mutant enzyme, where virtually no charge translocation corresponding to the A → P_R reaction was observed (see Supplementary Information), although electron transfer still occurs from haem *a* to the binuclear site¹⁵. The proton acceptor cannot be the binuclear site because proton uptake into this site would produce intermediate F, which is formed only in the next reaction. We suggest, therefore, that the charge translocation during A → P_R is due to internal transfer of a proton from E278 to a so far unidentified location above the haem groups^{2,14,20–22} (I → II, Fig. 3), which is the initial step of the proton pump mechanism. It is followed by protonation of E278 via the D-pathway (II → III), proton transfer to the binuclear site (III → IV), re-protonation of E278 and release of the pumped proton to the aqueous P-side. The latter reactions all occur secondarily during the P_R → F transition, which accounts for the fact that most of the charge translocation amplitude is kinetically linked to this reaction.

Our results are consistent with the recent observation² that proton release and uptake on the two aqueous sides of the membrane are kinetically coupled to the P_R → F reaction. However, we show here that these events are preceded by an internal vectorial proton transfer reaction that is kinetically associated with electron transfer from haem *a* to the O₂ reduction site, and which can be ascribed as the initial driving step of the proton pump mechanism (Fig. 1a, short blue arrow; Fig. 3, I → II). This finding has its parallel in the proton pump of bacteriorhodopsin, where mechanistically important internal proton transfers precede proton equilibration with the surroundings^{6,23}. Important details of the mechanism still remain unsolved, however, such as the identity of the proton-binding site above the haem groups, and the reason for why proton transfer from

E278 to that site occurs before transfer of the substrate proton to the binuclear site^{21,22,24}.

METHODS

Enzyme preparation and reconstitution into phospholipid vesicles. Bacterial growth, isolation of bacterial membranes and purification of cytochrome *c* oxidase from *P. denitrificans*, as well as reconstitution of wild-type and mutant enzymes into proteoliposomes, were done by described methods⁸, except that the enzyme concentration during reconstitution was increased to 6 μM.

Electrometric measurements. A detailed description of the electrometric method has been published⁸. The voltage recorded is proportional to the extent of charge translocation perpendicular to the membrane plane. The pH dependence of the electrometric response was measured in a variety of buffers: 100 mM MES (in the range pH 6.0–6.6), MOPS (pH 6.6–7.8), Tris (pH 7.8–8.8), Bis-Tris propane (pH 8.8–9.2), CHES (pH 9.2–9.8) and CAPS (pH 9.8–10.5). Media were supplemented with 50 mM glucose, 0.3 mg ml⁻¹ catalase, 3 mg ml⁻¹ glucose oxidase and 1 μM hexaammine ruthenium (III).

Transient optical spectroscopy. The reaction of the solubilized fully reduced wild-type enzyme with oxygen was measured on a microsecond to millisecond timescale. About 30 μM enzyme in 100 mM CAPS, pH 10.5, 0.02% dodecyl maltoside, was made anaerobic on a vacuum line, reduced with 20 mM potassium ascorbate and 5 μM TMPD (*N,N,N',N'*-tetramethyl-*p*-phenylenediamine) and incubated with 100% CO. The sample was loaded into one syringe of a stopped-flow system (RX2000, Rapid Kinetics Spectrometer Accessory, Applied Photophysics) and mixed with a fivefold volume of oxygen-saturated buffer. The reaction was started by a laser flash to photo-dissociate CO bound to the enzyme.

Received 22 September 2005; accepted 1 February 2006.

- Wikström, M. K. F. Proton pump coupled to cytochrome *c* oxidase in mitochondria. *Nature* **266**, 271–273 (1977).
- Faxén, K., Gilderson, G., Ådelroth, P. & Brzezinski, P. A mechanistic principle for proton pumping by cytochrome *c* oxidase. *Nature* **437**, 286–289 (2005).
- Babcock, G. T. & Wikström, M. O₂ activation and the conservation of energy in cell respiration. *Nature* **356**, 301–309 (1992).
- Ferguson-Miller, S. & Babcock, G. T. Heme/copper terminal oxidases. *Chem. Rev.* **96**, 2889–2907 (1996).
- Gennis, R. B. Coupled proton and electron transfer reactions in cytochrome oxidase. *Front. Biosci.* **9**, 581–591 (2004).
- Wikström, M. Proton translocation by bacteriorhodopsin and heme-copper oxidases. *Curr. Opin. Struct. Biol.* **8**, 480–488 (1998).
- Morgan, J. E., Verkhovskiy, M. I., Palmer, G. & Wikström, M. The role of the P_R intermediate in the reaction of cytochrome *c* oxidase with O₂. *Biochemistry* **40**, 6882–6892 (2001).
- Jasaitis, A., Verkhovskaya, M., Morgan, J. E., Verkhovskiy, M. & Wikström, M. Assignment and charge translocation stoichiometries of the major electrogenic phases in the reaction of cytochrome *c* oxidase with dioxygen. *Biochemistry* **38**, 2697–2706 (1999).
- Bloch, D. et al. The catalytic cycle of cytochrome *c* oxidase is not the sum of its two halves. *Proc. Natl Acad. Sci. USA* **101**, 529–533 (2004).
- Namslauer, A., Aagaard, A., Katsonouri, A. & Brzezinski, P. Intramolecular proton transfer reactions in a membrane-bound proton pump: the effect of pH on the peroxy to ferryl transition in cytochrome *c* oxidase. *Biochemistry* **42**, 1488–1498 (2003).
- Mitchell, R. & Rich, P. R. Proton uptake by cytochrome *c* oxidase on reduction and on ligand binding. *Biochim. Biophys. Acta* **1186**, 19–26 (1994).
- Ruitenbergh, M., Kannt, A., Bamberg, E., Fendler, K. & Michel, H. Reduction of cytochrome *c* oxidase by a second electron leads to proton translocation. *Nature* **417**, 99–102 (2002).
- Iwata, S., Ostermeier, C., Ludwig, B. & Michel, H. Structure at 2.8 Å resolution of cytochrome *c* oxidase from *Paracoccus denitrificans*. *Nature* **376**, 660–669 (1995).
- Brändén, G. et al. The protonation state of a heme propionate controls electron transfer in cytochrome *c* oxidase. *Biochemistry* **44**, 10466–10474 (2005).
- Ådelroth, P., Svensson-Ek, M., Mitchell, D. M., Gennis, R. B. & Brzezinski, P. Glutamate 286 in cytochrome *aa*₃ from *Rhodospirillum rubrum* is involved in proton uptake during the reaction of the fully reduced enzyme with dioxygen. *Biochemistry* **36**, 13824–13829 (1997).
- Thomas, J. W., Puustinen, A., Alben, J. O., Gennis, R. B. & Wikström, M. Substitution of asparagine for aspartate 135 in subunit I of the cytochrome *bo* ubiquinol oxidase of *Escherichia coli* eliminates proton pumping activity. *Biochemistry* **32**, 10923–10928 (1993).
- Babcock, G. T. How oxygen is activated and reduced in respiration. *Proc. Natl Acad. Sci. USA* **96**, 12971–12973 (1999).
- Verkhovskiy, M. I., Jasaitis, A., Verkhovskaya, M. L., Morgan, J. E. & Wikström, M. Proton translocation by cytochrome *c* oxidase. *Nature* **400**, 480–483 (1999).

19. Verkhovskaya, M. L. *et al.* Glutamic acid 286 in subunit I of cytochrome *bo*₃ is involved in proton translocation. *Proc. Natl Acad. Sci. USA* **94**, 10128–10131 (1997).
20. Michel, H. Cytochrome *c* oxidase: catalytic cycle and mechanism of proton pumping - a discussion. *Biochemistry* **38**, 15129–15140 (1999).
21. Wikström, M., Verkhovsky, M. I. & Hummer, G. Water-gated mechanism of proton translocation by cytochrome *c* oxidase. *Biochim. Biophys. Acta* **1604**, 61–65 (2003).
22. Popovic, D. M. & Stuchebrukhov, A. A. Proton pumping mechanism and catalytic cycle of cytochrome *c* oxidase: Coulomb pump model with kinetic gating. *FEBS Lett.* **566**, 126–130 (2004).
23. Lanyi, J. K. & Luecke, H. Bacteriorhodopsin. *Curr. Opin. Struct. Biol.* **11**, 415–419 (2001).
24. Siegbahn, P. E. M., Blomberg, M. R. A. & Blomberg, M. L. Theoretical study of the energetics of proton pumping and oxygen reduction in cytochrome oxidase. *J. Phys. Chem. B* **107**, 10946–10955 (2003).
25. Flöck, D. & Helms, V. Protein-protein docking of electron transfer complexes: cytochrome *c* oxidase and cytochrome *c*. *Proteins* **47**, 75–85 (2002).
26. Tsukihara, T. *et al.* The low spin heme of cytochrome *c* oxidase as the driving element of the proton-pumping process. *Proc. Natl Acad. Sci. USA* **100**, 15304–15309 (2003).
27. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
28. Yoshikawa, S. *et al.* Redox-coupled crystal structural changes in bovine heart cytochrome *c* oxidase. *Science* **280**, 1723–1729 (1998).
29. Buse, G., Soulimane, T., Dewor, M., Meyer, H. E. & Bluggel, M. Evidence for a copper-coordinated histidine-tyrosine cross-link in the active site of cytochrome oxidase. *Protein Sci.* **8**, 985–990 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Laakkonen for help in preparing Fig. 1, and A. Puustinen and C. Ribacka for providing samples of wild-type and mutant enzyme. This work was supported by grants from the Sigrid Jusélius Foundation, Biocentrum Helsinki and the Academy of Finland (programme 44895).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.W. (marten.wikstrom@helsinki.fi).

Crystal structure of the CorA Mg^{2+} transporter

Vladimir V. Lunin^{1*}, Elena Dobrovetsky^{2*}, Galina Khutoreskaya², Rongguang Zhang⁴, Andrzej Joachimiak⁴, Declan A. Doyle⁵, Alexey Bochkarev^{2,3,6}, Michael E. Maguire⁷, Aled M. Edwards^{1,2,3,4,6} & Christopher M. Koth^{2†}

The magnesium ion, Mg^{2+} , is essential for myriad biochemical processes and remains the only major biological ion whose transport mechanisms remain unknown. The CorA family of magnesium transporters is the primary Mg^{2+} uptake system of most prokaryotes^{1–3} and a functional homologue of the eukaryotic mitochondrial magnesium transporter⁴. Here we determine crystal structures of the full-length *Thermotoga maritima* CorA in an apparent closed state and its isolated cytoplasmic domain at 3.9 Å and 1.85 Å resolution, respectively. The transporter is a funnel-shaped homopentamer with two transmembrane helices per monomer. The channel is formed by an inner group of five helices and putatively gated by bulky hydrophobic residues. The large cytoplasmic domain forms a funnel whose wide mouth points into the cell and whose walls are formed by five long helices that are extensions of the transmembrane helices. The cytoplasmic neck of the pore is surrounded, on the outside of the funnel, by a ring of highly conserved positively charged residues. Two negatively charged helices in the cytoplasmic domain extend back towards the membrane on the outside of the funnel and about the ring of positive charge. An apparent Mg^{2+} ion was bound between monomers at a conserved site in the cytoplasmic domain, suggesting a mechanism to link gating of the pore to the intracellular concentration of Mg^{2+} .

The CorA magnesium transporter is a homopentamer with five-fold symmetry about a central pore and can be divided into three parts (Fig. 1). A carboxy-terminal transmembrane domain comprises two transmembrane helices from each monomer (Fig. 2). The middle portion resembles a funnel, narrow at the entrance (~5 Å) and wide at the mouth (~20 Å), that is formed largely by a long α -helix extension of the inner transmembrane helix. Finally, a large cytoplasmic domain lies exterior to the funnel.

The cytoplasmic domain of CorA is a seven-stranded parallel/antiparallel β -sheet ($\beta 2\Delta\beta 1\nabla\beta 3\Delta\beta 7\Delta\beta 6\nabla\beta 5\Delta\beta 4\nabla$) sandwiched between two sets of α -helices ($\alpha 1$, $\alpha 2$, $\alpha 3$) and ($\alpha 4$, $\alpha 5$, $\alpha 6$) (Fig. 1). The domain fold is unlike all other known structures of ion channels or transporters and constitutes a new protein fold (see Supplementary Information). This domain, solved in its soluble form at 1.85 Å resolution (Supplementary Fig. S1), is linked to the transmembrane helices by the long $\alpha 7$ helix (residues 251–312), termed the stalk helix. The stalk helix kinks as it enters the membrane, extends through the membrane, forms the first transmembrane helix (TM1; residues 293–312) and harbours the 'YGMNF' signature sequence of CorA (residues 311–315)^{5,6} (Fig. 2 and Supplementary Fig. S2). The five TM1 helices (residues 293–312) form the pore. After a short extracellular seven-amino-acid loop, the TM2 helix (residues 326–345) returns back to the cytoplasm and ends in a highly conserved C-terminal KKKKWL motif (Fig. 3). In the current

structure, neither the extracellular loop nor the final two amino acids could be resolved.

The cytoplasmic domain shows the lowest sequence conservation among CorA transporters. Nonetheless, the predicted secondary structure of CorA orthologues agrees well with the three-dimensional structure of the *T. maritima* CorA, allowing the prediction of the structure of both the prokaryotic ZntB Zn^{2+} efflux transporter⁷ and eukaryotic mitochondrial MRS2 homologues⁸ as CorA-like structures even though these proteins have minimal sequence conservation with CorA (Supplementary Fig. S3).

The transmembrane region comprises two concentric rings of helices. The TM2 helices, which are essentially perpendicular to the membrane, form an outer ring that envelops the inner ring of TM1 helices, which form the pore. The TM1 helices are twisted along the five-fold axis and kinked in two places, at Pro 303 and Gly 312 (Fig. 2), resulting in a tapering of the pore near the periplasmic surface. Pro 303 is conserved in a similar position in almost all CorA orthologues. The distortion of transmembrane helices by proline or glycine residues is a feature noted in other ion channels and can provide flexibility during the dynamic gating process^{9,10}.

The CorA structure appears to represent a closed form of the transporter. A series of constrictions are present along the transmembrane section of the ion conduction pathway. These start with a ring of five Asn 314 side chains at the periplasmic entrance that probably block ion movement (Figs 2 and 4). The Asn 314 ring appears to be held in conformation, at least in part, by an aromatic stacking interaction between Tyr 311 of one monomer and Phe 315 of an adjacent monomer (that is, Y and F of the conserved YGMNF sequence) (Fig. 2 and Supplementary Fig. S2). Immediately following this sequence is a loop connecting TM1 and TM2. The loop is poorly ordered in this structure (residues 316–322 are unassigned). Nevertheless, this sequence always contains significant negative charge in those CorA family members known to transport Mg^{2+} (Supplementary Fig. S4)^{6,11}. The presence of the negatively charged loop on the periplasmic face of these Mg^{2+} transporters strongly suggests that this region comprises part of a selectivity filter.

The cytosolic part of the transporter, shaped like a funnel, is formed by the amino-terminal domains. The five $\alpha 7$ stalk helices form both the wall of the pore (TM1) and the wall of the funnel (Fig. 1). The inner wall of the funnel is lined along its length with successive rings of negative or hydroxyl-bearing residues (Tyr 255, Asp 263, Glu 266, Asp 270, Ser 273, Asp 277, Ser 281, Ser 284, Glu 230, Ser 233, Asp 238, Glu 246) (Fig. 3). Such an arrangement of charged residues is found in other cation channels and may serve as an electrostatic sink to increase the local ion concentration¹².

The orientation of the outer TM2 helices positions the short, highly basic C terminus of the protein (KKKKWL) on the intracellular side

¹Department of Medical Biophysics, ²Banting and Best Department of Medical Research, ³Department of Medical Genetics and Microbiology, University of Toronto, 112 College Street, Toronto, Ontario M5G 1L6, Canada. ⁴Structural Biology Center and Midwest Center for Structural Genomics, Biosciences Division, Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, Illinois 60439, USA. ⁵Structural Genomics Consortium, Botnar Research Centre, Oxford, Oxon OX3 7LD, UK. ⁶Structural Genomics Consortium, Banting Institute, 100 College Street, Toronto, Ontario M5G 1L5, Canada. ⁷Department of Pharmacology, Case Western Reserve University, Cleveland, Ohio 44106-4965, USA. [†]Present address: Vertex Pharmaceuticals Incorporated, 130 Waverly Street, Cambridge, Massachusetts 02139, USA.

*These authors contributed equally to this work.

of the membrane. Given the pentameric nature of CorA, this places at least 20 lysines at the membrane interface, creating a ring of positive charge approximately 50 Å in diameter, which we term the basic sphincter (Fig. 3). Additional lysines from the $\alpha 7$ stalk helix (Lys 292 and Lys 286) may also contribute to this remarkable concentration of positive charge. Abutting the basic sphincter is a cluster of negatively charged residues from two helices in the soluble domain ($\alpha 5$ and $\alpha 6$; residues 166–202 and 208–238). These helices, which are oriented perpendicularly towards the membrane interface and ‘hang’ towards the membrane like the branches of a weeping willow tree, are referred to as the willow helices. Almost half the residues in the willow helices and their intervening loop are aspartate or glutamate, and are highly conserved (Fig. 3 and Supplementary Fig. S5).

An Mg^{2+} ion entering the cell through CorA would traverse a distance of at least 40 Å, extending through the putative periplasmic selectivity filter, transmembrane domain and part of the cytoplasmic domain (Fig. 4). The diameter of the pathway varies considerably along its length, producing two small cavities with diameters of ~ 6 Å and ~ 5.6 Å; these intersperse three constrictions. Starting from the periplasmic side, the first constriction is formed by the Asn 314 side chains. The second is formed by Thr 305 and Met 302, and the third, positioned at the membrane/cytoplasmic interface, is formed by two hydrophobic residues, Leu 294 and Met 291. There is experimental electron density inside the two cavities near Gly 309 and Ala 298 that is not readily assignable to backbone or side chain groups, consistent with the presence of Mg^{2+} ions. Potassium channels have a single, larger cavity, located at the

centre of the bilayer, which helps to stabilize an ion¹³. The two cavities possibly have an analogous role for CorA. The basic sphincter is at the same level in the protein as the last constricting hydrophobic girdle. This suggests that the concentration of positive charge could be important for controlling pore diameter at this point in TM1.

Given the varying diameter throughout the pore, it appears that the transporter is in a closed state. The hydrophobic ring created by residues Leu 294 and Met 291 would create a strong energetic barrier to ion permeation and therefore be expected to form the main gate. A bulky hydrophobic residue at position 294 is a conserved feature of CorA transporters. This hydrophobic gate seems to be one of a number of recurring themes in cation selective channels^{12,14} by which ion movement is controlled indirectly through the controlled movement of water. This has been observed in the closed-state structures of the potassium channel KirBac1.1¹⁵, the mechanosensitive stretch receptor MscL¹⁶ and the nicotinic acetylcholine receptor (nAChR)¹⁷.

Even though CorA has several structural themes in common with ion channels it transports ions much more slowly. The nAChR allows 100,000 ions per second through whereas K^+ channels allow up to 100 million ions through per second (ref. 18). The throughput of CorA is only of the order of 10 million ions per minute for the whole cell³. This must be related to the way in which the Mg^{2+} ion interacts with the transporter. The problem may be that Mg^{2+} prefers to be hexacoordinated and probably remains so throughout the conduction process¹⁹. Hence, as the hydrated Mg^{2+} ion enters the conduction pathway, the number of available coordination sites would not match the five-fold symmetry of the CorA selectivity filter. The key to

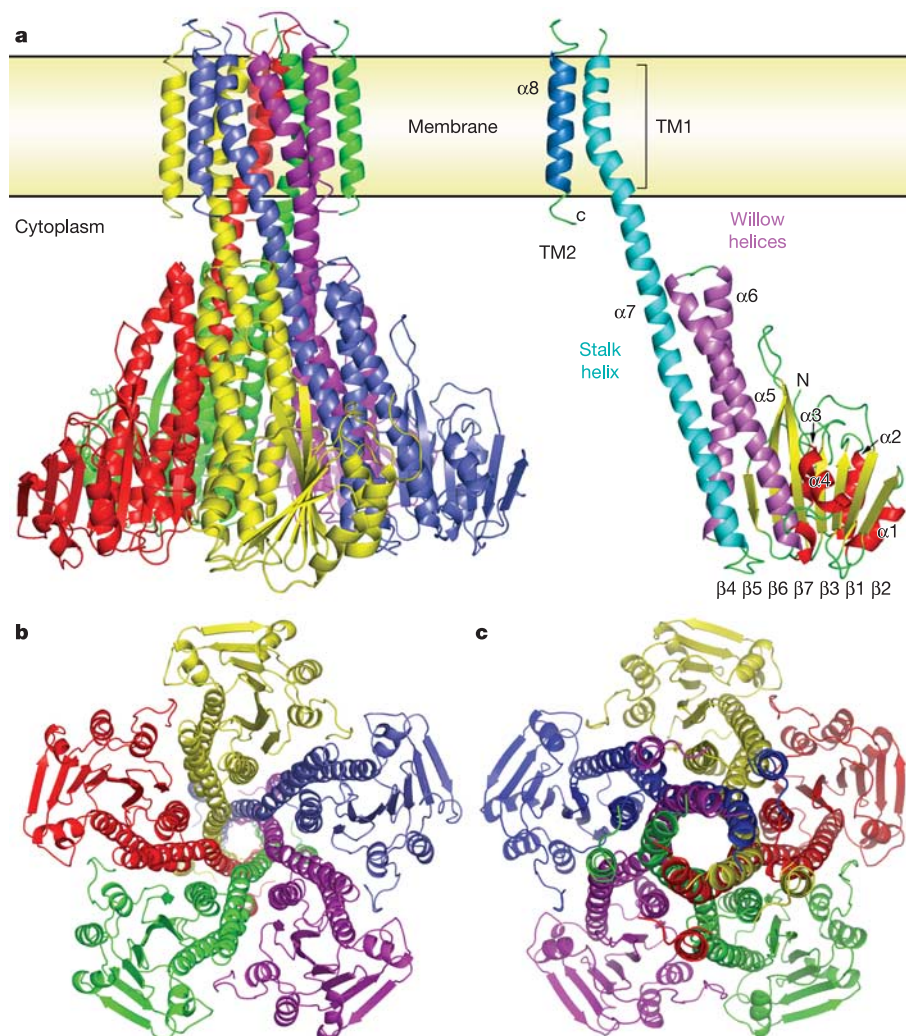
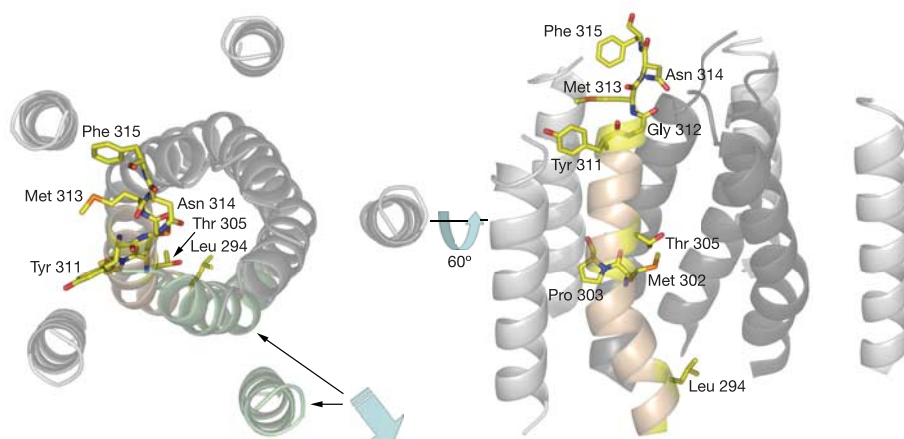


Figure 1 | Structure of the CorA Mg^{2+} channel.

a, Ribbon diagram of the CorA pentameric complex, viewed in the plane of the membrane. On the right is a single unit of the CorA channel highlighting the following structural features: Stalk helix and inner TM1 helix (turquoise), outer helix (dark blue) and willow helices (purple). These and the remaining α -helices (red) and β -sheets (yellow) are numbered according to the text. **b**, View from the intracellular region. **c**, View from the periplasm. This and other figures were generated with the program PyMOL²⁵, except where indicated.

**Figure 2 | Analysis of the CorA pore.**

The CorA pore, highlighting the locations of key residues on one of the TM1 helices (pale brown), as discussed in the text. The view is looking down the pore from the periplasmic surface (left panel) and in the plane of the membrane (right panel, orientation in left panel rotated 60° towards the viewer). To illustrate the pore interior better, a pair of TM1 and TM2 helices was removed (highlighted in pale green and indicated by blue arrow) for the panel on the right.

K⁺ channel selectivity and throughput is the high degree to which the potassium channel's selectivity filter mimics the hydration coordination of potassium ions^{13,14}. Hence, the hydrated Mg²⁺/CorA symmetry mismatch may be a form of control that slows down the rate at which the Mg²⁺ ion can move through CorA.

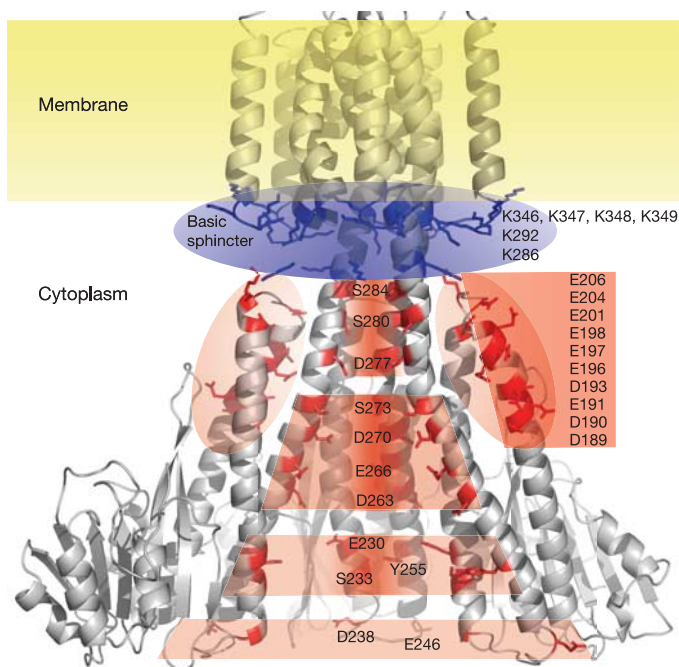
The selectivity determinants of *T. maritima* CorA for Mg²⁺ are unclear because part of the periplasmic entrance to the pore, which presumably contains an ion selectivity filter, was not seen in the structure. However, the visible structural features of the CorA transporter, as well as the unique chemistry of Mg²⁺, limit the mechanisms by which it can be highly selective for the ion. Sieving for hydrated cations is an insufficient explanation because the hydrated radius of magnesium is the largest of all biological cations (~5 Å)¹⁹. Based on size alone, Ca²⁺ would be predicted to be transported, but it is neither a substrate nor an inhibitor of CorA^{3,20}. Moreover, cobalt(III)hexaammine, which has the same size as a hydrated Mg²⁺ but carries a +3 charge, potentially inhibits transport via CorA transporters²¹. Thus, the CorA transporter must make use of other properties of ions, in addition to their charge and hydrated size.

For several reasons, we believe that an incoming magnesium ion is at least partially dehydrated upon entering the ion conduction pore of CorA, and that this is a key basis for selectivity. First, CorA transporters can also transport Co²⁺ and Ni²⁺, two ions with atomic radii similar to that of Mg²⁺, but not Mn²⁺, whose atomic radius is significantly larger than Mg²⁺ (refs 3, 20). Second, the atomic Mg²⁺ ion is the smallest of all biological cations, ~350 times smaller than its hydrated form¹⁹. A conduction pathway that exploited the small size of the dehydrated Mg²⁺ ion would conceivably be highly selective. Third, the inhibitory action of cobalt(III)hexaammine, which is almost exactly the same size as a hydrated Mg²⁺, suggests that a fully hydrated Mg²⁺ ion first binds to the CorA transporter but cannot enter the pore. Fourth, the lining of the periplasmic entrance to the pore with polar groups (Figs 2, 4 and Supplementary Fig. S2) may serve to compensate for the energetic cost of dehydration. The Asn 314 constricting ring, which obviously cannot block the pore in an open conformation, may have a role in the initial Mg²⁺ dehydration, as occurs in other ion channels^{12,14}.

The crystal structure of the isolated cytoplasmic domain at 1.85 Å resolution revealed electron density hexacoordinated to oxygen atoms (one from Asp 89 and five from water molecules) that can be reasonably interpreted as a Mg²⁺ ion (Fig. 5a and see Supplementary Information). The structure of the full-length CorA, although at lower resolution, revealed strong electron density at a similar location, between the cytoplasmic domains of adjacent monomers (Fig. 5b). This density was found proximal to residue Asp 89 in the α3 helix of one monomer and Asp 253 on the stalk helix (α7) of the adjacent monomer. Asp 89 and Asp 253 are among the most highly conserved residues, pointing to a critical functional role. We suggest

that this putative Mg²⁺-binding site could be the pivot point of a hinge comprising the stalk helix and the α-β-α sandwich domain, allowing metal binding to influence the orientation of these elements within the CorA pentamer. Small changes in the structure of the Mg²⁺-binding site might be predicted to translate into large movements of both the stalk helix and the willow helices at their opposite ends near the pore.

How might the binding of Mg²⁺ be exploited for regulation? CorA must have a mechanism to regulate pore opening in response to changes in the intracellular concentration of ion. It is noteworthy that the narrowest part of the pore, the putative Leu 294/Met 291 gate, is precisely at the level of the basic sphincter noted above. The sphincter might be predicted to draw negative charge away from the middle of the pore at this level, preventing passage of the positively charged Mg²⁺ cation. The presence of a putative Mg²⁺-binding site

**Figure 3 | Electrostatic view of CorA.** Positively charged residues are coloured blue and negatively charged or hydroxyl-containing residues are coloured red. Charged residues within the CorA basic sphincter (blue-highlighted region) and cytoplasmic domain (funnel interior and willow helices, red-highlighted regions) are labelled and illustrated as stick models. For better viewing of the funnel interior, two of the CorA monomers were removed from the model.

in the soluble domain links important structural elements within the transporter and suggests a mechanism to open the gate. In the presence of sufficient magnesium, the α - β - α domain would be kept near to the stalk helix of the adjacent monomer, pulling the willow helices away from the membrane and the basic sphincter residues apparently closing the pore. Loss of Mg^{2+} from the binding site between $\alpha 7$ of one monomer and $\alpha 3$ of the adjacent monomer would allow rotation or other movement of the α - β - α domain. This would necessarily cause movement of the $\alpha 7$ stalk helix and of the acidic willow helices relative to the basic sphincter. Because the membrane-proximal portions of the willow helices are highly negatively charged, this would in turn influence the effective concentration of positive charge created by the basic sphincter, perhaps causing it to relax. Coupled with movement of the extremely long $\alpha 7$ helix, this would allow the gate formed by Leu 294/Met 291 to open, allowing Mg^{2+} through.

Key residues within the pore-forming TM1 domain can be interpreted with regards to *Salmonella enterica* serovar Typhimurium mutagenesis data²². Leu 294 appears to block the pore near the cytoplasmic face. The equivalent residue in *S. Typhimurium* is Phe 259, an equally bulky residue. Mutation of the adjacent Ser 260 residue in *S. Typhimurium* shows that substitution of smaller (Ala) but not larger (Val) residues allows transport, a result compatible with Leu 294/Phe 259 forming a moveable barrier for transport. Mutation of Pro 268 and Pro 269 in *S. Typhimurium* reduces transport, compatible with Pro 303 in *T. maritima* forming a kink and functional hinge in the TM1 helix. Mutations within the signature YGMNF sequence (residues 311–315 in *T. maritima* and 276–280 in *S. Typhimurium*) almost invariably abolish transport. This suggests

that this sequence is critical both for positioning the periplasmic loop that is presumably responsible for initial binding of Mg^{2+} , and that Asn 314 is possibly required to restrict the external entrance to the pore. Moreover, of the multiple mutations made in this sequence in *S. Typhimurium*, Tyr 311 is the only residue that can be changed with retention of activity. In the structure of CorA, the side chain of Tyr 311 is pointed away from the pore, and appears to form an aromatic stacking interaction with Phe 315 of an adjacent monomer. Mutation of Tyr 311 to tryptophan, which would probably retain this interaction, has only modest effects on activity.

We have crystallized the *T. maritima* CorA Mg^{2+} transporter in an apparently closed state. Like the MscL mechanosensitive ion channel¹⁶, glycine²³, acetylcholine¹⁷, GABA_A and serotonin (5HT₃) receptors²⁴, the CorA transporter is assembled as a pentamer. However, beyond this common pentameric organization, the CorA

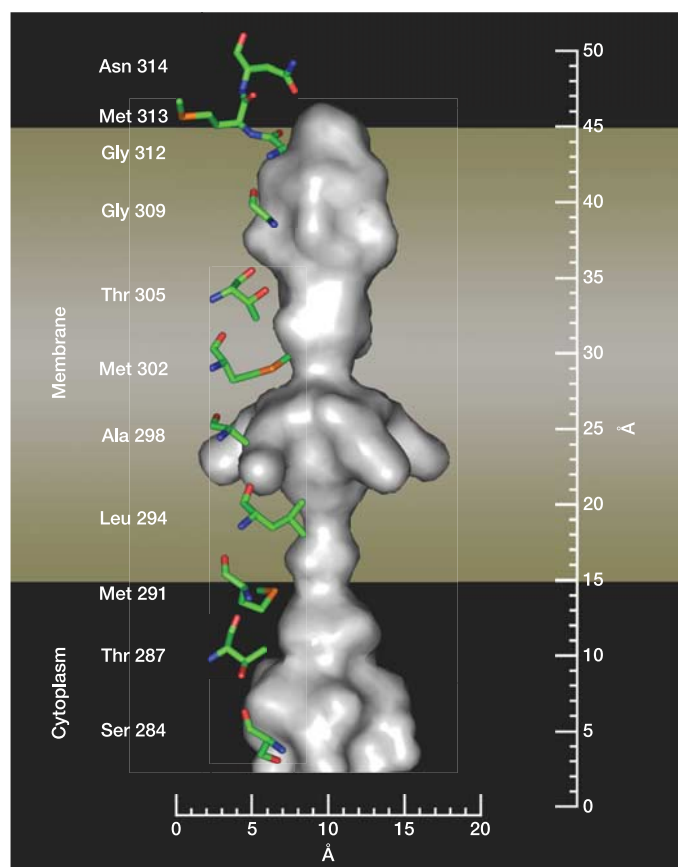


Figure 4 | Pore dimensions. A cut-away view displaying the solvent-accessible surface and dimensions of the CorA pore. Also shown are stick models of residues lining the pore. The key residues discussed in the text are labelled.

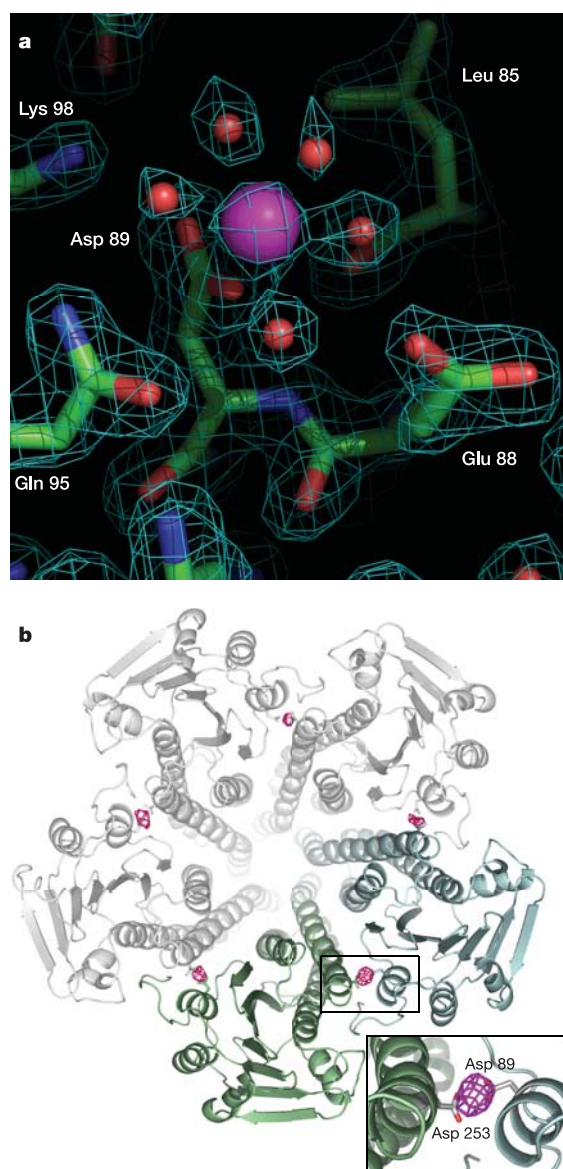


Figure 5 | Electron density corresponding to protein and bound magnesium. **a**, A portion of the soluble domain 1.85 Å electron density map in the region of Asp 89, showing electron density for the putative Mg^{2+} ion and the water molecules that fill the hexacoordination shell. **b**, A portion of the 3.9 Å difference Fourier electron density map (purple) showing the putative magnesium ion between monomers. The inset shows the residue Asp 89 in one monomer (pale blue) and Asp 253 in another monomer (green) and a peak in the difference Fourier electron density map.

transporter is quite different. On the basis of this structure, we propose that several key principles underlie the mechanism of Mg^{2+} transport through CorA. Initial selectivity occurs at the periplasmic surface and involves dehydration. In the absence of sufficient intracellular Mg^{2+} levels, Mg^{2+} ions bound between monomers are released. This causes the willow helices to shift and the basic sphincter, which is in register with the Leu294/Met291 gate, to open and allow Mg^{2+} ions to flow through. In addition to the Leu294/Met291 gate and the basic sphincter, other regions of CorA probably widen to permit ion transport. In short, selectivity of CorA transporters is achieved at the periplasmic face, but intracellular Mg^{2+} determines transport. Clearly, support for this mechanism will require further functional studies. Nevertheless, the structure of *T. maritima* CorA should serve as an attractive model for understanding the mechanisms regulating Mg^{2+} ion transport in this broad protein family.

METHODS

Expression of full-length and soluble domain of *T. maritima* CorA. The coding sequence of *T. maritima* CorA and its N-terminal domain (residues 1–266) were expressed in *Escherichia coli* as fusions to an N-terminal six-histidine tag in a modified T7 polymerase expression vector. After growth of the cells and induction of protein expression, the cells were harvested, resuspended, lysed and the membrane fraction isolated. CorA was solubilized from this fraction in a solution containing 50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 10 mM imidazole, 1% n-dodecyl- β -D-maltopyranoside (DDM, Anatrace) and protease inhibitors and then purified by nickel affinity chromatography. Expression of selenomethionine-labelled proteins was essentially as described²⁵.

Crystals of full-length *T. maritima* CorA and the CorA soluble domain were grown in hanging drops over a solution containing either 20% (w/v) PEG 2000 (Fluka), 0.3 M $Mg(NO_3)_2$ and 0.1 M Tris pH 8.0, or 35% (w/v) PEG 3350 (Fluka), 0.2 M $MgCl_2$ and 0.1 M Tris pH 8.5, respectively. Detailed purification and crystallization procedures can be found at <http://sgc.utoronto.ca/SGC-WebPages/StructureDescription/2BB1.php> and in the Supplementary Information.

Structure determination. Native crystals of full-length CorA diffracted X-rays to 3.9 Å. Selenomethionine-containing crystals formed only from protein preparations that had low incorporation (<20%) of selenomethionine. The low occupancy prevented us from solving the selenium substructure from these data. Phase information for the crystals of the full-length protein was obtained by first determining the 1.85 Å structure of the amino-terminal two-thirds of CorA (residues 1–244) and using this structure as a search model for molecular replacement.

The final model of the N-terminal domain of CorA contained residues 13–117, 120–199 and 207–244, 236 water molecules and two metal ions, assigned as Mg^{2+} and Na^+ , and four molecules of detergent (DDM). The model for the full-length protein, derived from the molecular replacement solution, was built initially by fitting the N-terminal domain into the electron density and then improved by a combination of building, refining and five-fold non-crystallographic averaging. To help to register the amino acid sequence within the mostly helical electron density, phases taken from this improved model were used to calculate an anomalous Fourier map from the data collected from selenomethionine-containing crystals. Inspection of this map revealed eight peaks and their non-crystallographic symmetry mates; three peaks corresponded exactly to the positions of the Met residues in the soluble domain. There were two additional peaks inside the pore and three outside the pore. The register of the assigned sequence based on the conveniently placed methionine signposts corresponds well with two kinks in the stalk helix at positions Gly274 and Pro303. The final model, which was refined to 3.9 Å, contains residues 9–315 and 323–349. The *R* factor was 0.361, and *R*_{free} = 0.406.

Received 27 October 2005; accepted 9 February 2006.

- Nelson, D. L. & Kennedy, E. P. Magnesium transport in *Escherichia coli*. Inhibition by cobaltous ion. *J. Biol. Chem.* **246**, 3042–3049 (1971).

- Hmiel, S. P., Snively, M. D., Miller, C. G. & Maguire, M. E. Magnesium transport in *Salmonella typhimurium*: Characterization of magnesium influx and cloning of a transport gene. *J. Bacteriol.* **168**, 1444–1450 (1986).
- Bui, D. M., Gregan, J., Jarosch, E., Ragnini, A. & Schwenen, R. J. The bacterial magnesium transporter CorA can functionally substitute for its putative homologue Mrs2p in the yeast inner mitochondrial membrane. *J. Biol. Chem.* **274**, 20438–20443 (1999).
- Kehres, D. G. & Maguire, M. E. Structure, properties and regulation of magnesium transport proteins. *Biomaterials* **15**, 261–270 (2002).
- Kehres, D. G., Lawyer, C. H. & Maguire, M. E. The CorA magnesium transporter gene family. *Microb. Compar. Genom.* **43**, 151–169 (1998).
- Worlock, A. J. & Smith, R. L. ZntB is a novel Zn^{2+} transporter in *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* **184**, 4369–4373 (2002).
- Gardner, R. C. Genes for magnesium transport. *Curr. Opin. Plant Biol.* **6**, 263–267 (2003).
- Cordes, F. S., Bright, J. N. & Sansom, M. S. Proline-induced distortions of transmembrane helices. *J. Mol. Biol.* **323**, 951–960 (2002).
- Tieleman, D. P., Shrivastava, I. H., Ulmschneider, M. R. & Sansom, M. S. Proline-induced hinges in transmembrane helices: possible roles in ion channel gating. *Proteins* **44**, 63–72 (2001).
- Knoop, V., Groth-Malonek, M., Gebert, M., Eifler, K. & Weyand, K. Transport of magnesium and other divalent cations: evolution of the 2-TM-GxN proteins in the MIT superfamily. *Mol. Genet. Genom.* **274**, 205–216 (2005).
- Doyle, D. A. Molecular insights into ion channel function. *Mol. Membr. Biol.* **21**, 221–225 (2004).
- Roux, B. & MacKinnon, R. The cavity and pore helices in the KcsA K^+ channel: electrostatic stabilization of monovalent cations. *Science* **285**, 100–102 (1999).
- Doyle, D. A. *et al.* The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. *Science* **280**, 69–77 (1998).
- Kuo, A. *et al.* Crystal structure of the potassium channel KirBac1.1 in the closed state. *Science* **300**, 1922–1926 (2003).
- Chang, G., Spencer, R. H., Lee, A. T., Barclay, M. T. & Rees, D. C. Structure of the MscL homolog from *Mycobacterium tuberculosis*: A gated mechanosensitive ion channel. *Science* **282**, 2220–2226 (1998).
- Miyazawa, A., Fujiyoshi, Y. & Unwin, N. Structure and gating mechanism of the acetylcholine receptor pore. *Nature* **423**, 949–955 (2003).
- MacKinnon, R. Potassium channels and the atomic basis for selective ion conduction. *Biosci. Rep.* **24**, 75–100 (2004).
- Maguire, M. E. & Cowan, J. A. Mg^{2+} chemistry and biochemistry. *Biomaterials* **15**, 203–210 (2002).
- Snively, M. D., Florer, J. B., Miller, C. G. & Maguire, M. E. Magnesium transport in *Salmonella typhimurium*: $^{28}Mg^{2+}$ transport by the CorA, MgtA, and MgtB systems. *J. Bacteriol.* **171**, 4761–4766 (1989).
- Kucharski, L. M., Lubbe, W. J. & Maguire, M. E. Cation hexaamines are selective and potent inhibitors of the CorA magnesium transport system. *J. Biol. Chem.* **275**, 16767–16773 (2000).
- Szegedy, M. A. & Maguire, M. E. The CorA Mg^{2+} transport protein of *Salmonella typhimurium*. Mutagenesis of conserved residues in the second membrane domain. *J. Biol. Chem.* **274**, 36973–36979 (1999).
- Langosch, D., Thomas, L. & Betz, H. Conserved quaternary structure of ligand-gated ion channels: the postsynaptic glycine receptor is a pentamer. *Proc. Natl Acad. Sci. USA* **85**, 7394–7398 (1988).
- Lester, H. A. The permeation pathway of neurotransmitter-gated ion channels. *Annu. Rev. Biophys. Biomol. Struct.* **21**, 267–292 (1992).
- Delano, W. L. PyMOL (<http://pymol.sourceforge.net/>) (cited 19 October 2005).
- Korolev, S. *et al.* The crystal structure of spermidine synthase with a multisubstrate adduct inhibitor. *Nature Struct. Biol.* **9**, 27–31 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The research was supported by grants from Genome Canada, the National Institutes of Health, the US Department of Energy, Office of Biological and Environmental Research, and the Structural Genomics Consortium. We thank M. McMillan and Y. Kim for assistance in collecting data, and A. Savchenko and the Indians and Red Wings screening teams for their guidance. We also thank C. Charky and D. Bouchard at Nextal Biotechnologies for their assistance throughout our membrane protein crystallization efforts.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.M.E. (aled.edwards@utoronto.ca) or M.E.M. (mem6@cwrw.edu).

-  FOCUS
-  SPOTLIGHT
-  RECRUITMENT
-  ANNOUNCEMENTS
-  EVENTS

naturejobs

The long road to equality

A recent report from the Institute of Physics (IOP), *Women in University Physics Departments*, details a survey during which researchers visited the physics departments of UK and Irish universities. They were assessing how female-friendly these departments are, and the results make interesting reading (www.iop.org). But as I was perusing the report, I found myself reminded of an advertising campaign for a cigarette aimed specifically at women.

Why? Well, the campaign debuted in 1968 with adverts that ran under the tagline: "You've come a long way, baby". This line contains anachronisms and contradictions that can be related to both women smokers of the late 1960s and women physicists today. The adverts purported to celebrate women's progress, but by using the word "baby" they instead denigrate it. And the site visits, which were aimed at improving the lot of women physicists, instead show just how little progress has been made.

Volunteers from the IOP went to about 40% of all physics departments in Britain and Ireland. Although they did see some signs that women physicists have come "a long way" — the number of women undergraduates has increased, for

example — they also saw signs that the sexist, unspoken "baby" still pervades. These signs included offensive posters in some labs, few, if any, women speakers at colloquia and seminars, and little formal efforts to recruit women physicists.

This might point to why women physicists are not well represented at the higher levels — especially compared with other disciplines such as the life sciences. None of the labs visited had female department heads and only 4% of tenured physics professors in Britain and Ireland are women. Such under-representation, perhaps combined with institutional sexism, might give young women scientists pause in pursuing physics as a career.

The IOP report carries its own tagline: "Creating female-friendly environments". Perhaps there should be a second note saying: "You've got a long way to go..."



Paul Smaglik, Naturejobs editor

CONTACTS

Publisher: Ben Crowe
Editor: Paul Smaglik
Assistant Editor: Corie Lok

European Head Office, London
 The Macmillan Building,
 4 Crinan Street
 London N1 9XW, UK
 Tel: +44 (0) 20 7843 4961
 Fax: +44 (0) 20 7843 4996
 e-mail: naturejobs@nature.com

European Sales Manager:
 Andy Douglas (4975)
Natureevents: Silke Opstrup (4994)

UK/RoW/Ireland/Italy:
 Nils Moeller (4953)
 Irene Viglia-Atton (4944)
Scandinavia/Spain/Portugal:
 Evelina Rubio Håkansson (4973)
France/Switzerland/Belgium:
 Amelie Pequignot (4974)
**Germany/Austria/
 The Netherlands:**
 Reya Silao (4970)

Advertising Production Manager:
 Billie Franklin
 To send materials use London
 address above.
 Tel: +44 (0) 20 7843 4814

Fax: +44 (0) 20 7843 4996
 e-mail: naturejobs@nature.com
Naturejobs web development: Tom Hancock
Naturejobs online production: Niamh Shields

**European Satellite Office
 Germany/Austria/
 The Netherlands:**
 Patrick Phelan
 Tel: +49 89 54 90 57 11
 Fax: +49 89 54 90 57 20
 e-mail: p.phelan@nature.com

US Head Office, New York
 75 Varick Street,
 9th Floor, New York,

NY 10013-1917
 Tel: +1 800 989 7718
 Fax: +1 800 989 7103
 e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

Japan Head Office, Tokyo
 Chiyoda Building,
 2-37 Ichigayatamachi,
 Shinjuku-ku,
 Tokyo 162-0843
 Tel: +81 3 3267 8751
 Fax: +81 3 3267 8746
Asia-Pacific Sales Director: Rinoko Asami
 e-mail: rasami@naturejpn.com

Breaking into business

T. GARCHA/ZEFA/CORBIS



While attending a lecture by a professor who had founded a drug company, Peter Kolchinsky had an idea that rerouted his career. After the talk, Kolchinsky helped found the Harvard Biotech Club, to let students and postdocs learn more about the industry. Every few weeks, the club hosted a different speaker. "By the time I graduated," Kolchinsky recalls, "I had met dozens of executives, venture capitalists, attorneys and entrepreneurs." Instead of going on to a postdoc after completing his PhD in virology, Kolchinsky co-founded RA Capital Management, a biotech investment company.

Although few scientists have started their own hedge funds, many carve business niches by drawing on their scientific training. There are many routes from the lab bench to the business world. Consulting firms recruit PhD students straight from campus, and university licensing offices are keen to put students' scientific expertise to work. Industrial scientists team up with business-development groups at their own companies. And many opt for formal training with an MBA.

No matter what route they choose, scientists face a common barrier: presenting themselves not as scientists, but as potential businesspeople. "Most scientists trying to break into business come across the same: desperate and unqualified," says Kolchinsky. That's because they seem more interested in running away from science than in pursuing business. To avoid

Forget what your mother told you. Scientists looking for jobs in the business world need to learn to talk to strangers, says Monya Baker.

that, scientists need to study the language and mindset of the business world.

Students can start preparing for business careers during their PhD. Kolchinsky advises them to pursue topics relevant to biotech, such as medicinal chemistry, pharmacology or specific disease research. Although the scientific knowledge most valuable for the business world tends to come early in graduate school, sticking it out can mean more opportunities later. For those with PhDs, proving scientific credentials can be as easy as flashing a business card, says Kolchinsky, "so don't drop out unless you're desperate".

Taking an interest

Graduate students don't need formal training to start exploring. Angie You began her PhD at Harvard University fully intending to become a professor, but something didn't feel right. So she started looking for other options. She entered student business-plan competitions, an increasingly common way to break into industry. She realized that she liked being grilled by venture capitalists and enjoyed reading business periodicals.

When You graduated, she began what she calls "a postdoc in business" with consulting firm McKinsey in San Francisco. She worked about the same hours as in the lab, but with more deadlines and travel. Consultants often set up camp with clients, interview scores of employees and outside experts, pore over

company and industry data, then make recommendations about business strategies, such as whether a company should develop drugs for particular markets. Connections she made while at McKinsey led You to several positions; she's now a vice-president at Venrock Associates, a venture-capital firm, also in northern California.

Killu Tougu Sanborn took a similar route, but through the technology-licensing office at the Memorial Sloan-Kettering Cancer Center in New York, rather than a consultancy. Volunteering there showed her how science was applied outside academic settings and taught her to evaluate new technologies. Sanborn planned to work at the licensing office after finishing her PhD to save money for business or law school. But before that happened, she attended a scientific conference where she met an executive from research-products company Stratagene in San Diego. He advised her that the best way to learn about business was through sales, another growing strategy for scientists to get a foot in the business door. Although the executive told Sanborn that Stratagene didn't hire people without experience, she applied anyway and got the job. She now heads the life-science venture-lending practice of Comerica Bank in San Diego.

A toe in the water

For industrial scientists who haven't yet left the bench, the safest place to try something else may be within their own companies. Zak Zimmerman was always interested in business, but after completing his postdoc, he took a scientist position at 454 Life Sciences in Branford, Connecticut, a 50-person start-up, because he liked the technology and wanted industry experience. When the business development team there wanted a scientist to explain current applications of the technology, Zimmerman not only did so but helped identify potential customers.

At a conference, Zimmerman met executives preparing to launch a healthcare practice at IDC, a large consulting firm based in Framingham, Massachusetts. They were impressed to find a scientist who had done market analysis for disruptive technologies, and Zimmerman thought the consultancy was a good vantage point from which to explore the business world. And so it turned out: Zimmerman is now director of external alliances at Alnylam Pharmaceuticals, but he says that anyone can make connections, particularly with conference speakers.

Other scientists turn to business after becoming frustrated by the lack of money or power in research departments. Andrew McDonald says he helped to invent one of the drugs that allowed his employer, Cytokinetics in South San Francisco, California, to go public, only to realize that scientific contributions wouldn't make him rich. In frustration, he began going to business school at the University of California, Berkeley, part-time. A year later, he quit. "I realized that the MBA wasn't necessary," he says, "but what was really valuable was my knowledge of drug discovery." He started to do interviews to find out what jobs would be right for him. After being introduced to the chief executive of investment bank ThinkEquity Partners by a casual acquaintance, McDonald was offered a position analysing life-sciences stocks.

Business employers have become increasingly open

HOTDOGS AND HOT DEALS

Shortly after Eric Gordon announced his retirement from the drug industry, an old acquaintance took him to a baseball game. Gordon had met John Freund in the 1990s, while pitching a company to venture capitalists. Freund wasn't interested, but the two kept in touch, and Freund went on to found Skyline Ventures. In 2002, over beer and hot dogs, Freund asked Gordon to help evaluate companies' technologies. Now, he is a partner at Skyline Ventures, where he works about half-time. The arrangement suits him because he still considers himself a scientist first; he finds it hard to get very interested in "such sacred things as a term sheet", he says. But that hasn't been a problem, he says. "I don't have to do it. I'm in a team."

M.B.

to hiring people who are moving from non-business careers, says Jürg Eckhardt, who manages a biotech venture fund for Swissfirst Asset Management. In the mid-1990s, Eckhardt became the first physician hired by McKinsey as a consultant in Switzerland. Despite this foundation of consulting, he still pursued an MBA. "After business school you're not seen as a scientist who is doing something in business, but as a broadly educated businessperson," Eckhardt says.

Those with the deepest research experience may have the most difficulty making the transition. "Business people think scientists waste time on being perfect instead of practical," says Kolchinsky.

That was precisely the attitude faced by Manish Singh, now a director at California Technology Ventures in Pasadena. He held scientific and managerial positions at companies including Genetic Therapy and Viagene. But after those two were acquired by Novartis and Chiron, respectively, he realized that business development teams, not industry scientists, were the ones who had the real say in what experiments were run. Yet even when he talked to executives about business development jobs, he would still receive offers for research positions.

On course

So Singh entered an MBA programme at the University of California, Los Angeles. Despite tuition fees and two years' lost salary, the programme was worth it, says Singh. He compares an MBA programme to a courtyard full of doors: "You get to open all the doors and see if that is the room that you really want to stay in." Singh looked in on hedge funds and various forms of business development and ultimately entered venture capital, taking an internship at the firm where he now has a permanent job.

Still, those who have made the transition say many scientists won't be happy in business. With an increased emphasis on milestones and deadlines, business people don't have time for the kind of deep expertise that scientists are accustomed to. The shift from the self-reliance that comes with planning one's own experiments to the team-building necessary to get deals done can also be discomfiting, particularly for academic scientists. Soft skills are necessary, not just for getting the job done, but also for getting the next job. After all, the best positions aren't advertised. People recruit through their networks.

And that harks back to the most important skill of all: meeting new people and making them your allies. "You have to be someone who the other person wants to have dinner with," says Singh. "You don't do business with people you don't like."

Monya Baker is a freelance writer based in San Francisco.



Zak Zimmerman found his way into business with a little help from networking at conferences.

MOVERS

Fotis Kafatos, chairman, scientific council of the European Research Council



2005–present: Chair, insect immunogenomics, Imperial College London

1993–2005: Director-general, EMBL, Heidelberg, Germany

1982–present: Professor of biology, University of Crete, Greece

1972–82: Professor of biology, Athens University, Greece

1969–94: Professor of biology, Harvard University

Being the leader of the scientific council of the European Research Council (ERC) when it launches later this year will give biologist Fotis Kafatos the chance to pass on a favour.

One of the ERC's aims is to support independent careers for early-stage investigators; Kafatos credits mentoring by an early adviser, Tom Eisner, a Cornell University chemical ecologist, for helping him become an expert in the molecular study of developmental biology and comparative genomics.

Kafatos, who survived malaria as a child, went on to lead the consortium that sequenced the mosquito genome in 2002 and continues to work on malaria.

"A lot of what I became was due to Tom's vision of how evolution permeates all of biology," he says. "He also taught me how important it is to be a mentor to others. You give and you receive, as an exchange not with one person but with a community."

The other goals of the ERC also resonate with Kafatos and his career. One of the ERC's initial programmes is to fund established researchers' work at the interface between fields. As a young professor at Harvard University in Cambridge, Massachusetts, he knew that advancements would be made with interdisciplinary approaches.

"I wanted to do research in insect metamorphosis," he says. "But I realized the field would only be cracked with the help of molecular biology, and that wasn't the vision of the time. I learnt molecular biology late, but it paid off."

Kafatos hopes the ERC will help keep good researchers in Europe and lure them back from abroad, as a part-time job at the University of Athens did for him. For many years, he divided his time between there and his tenured post at Harvard. He went on to build a department and research institute in his native Crete and then to lead the European Molecular Biology Laboratory (EMBL) in Germany.

When his term at EMBL ended last year, he moved to Imperial College London. He also turned 65, but that didn't mean a lot to him. "After my so-called 'retirement', I couldn't wait to get back into full-time research," he says.

Yet, even though it would require more administrative duties, the ERC job was too good to turn down. "It was a personal challenge and an opportunity to contribute to a project of vital importance for science in Europe," he says. Luckily, it will still allow for more research time than he had at EMBL, where he went on to the lab after a working day.

"I believe passionately in the ERC," he says. "That's why I'm giving it more time than I really have!" ■

Janet Wright

SCIENTISTS & SOCIETIES

A meeting of biomedical minds

Science is truly an international enterprise. Never was this more apparent than at the World Life Sciences Forum BioVision in Lyon, France, last April. Occurring once every two years, this meeting brings together scientists, businesspeople and policy-makers for discussion and debate on the top issues in life sciences.

Attending that conference was one of the most memorable experiences of my young scientific career. Not only did I enjoy talks by world-renowned speakers on topics ranging from agricultural sciences to vaccine development, but I also got to meet and talk to several Nobel laureates and become inspired by their remarks.

But the most mind-blowing part was being one of the 100 young PhDs and MBAs from around the world chosen to participate in the conference as BioVision.Nxt fellows. During a day-long meeting before the conference, we had stimulating discussions on current issues such as the role of pharmaceutical companies and non-governmental organizations in research and the growing global threat of bioterrorism. It was wonderful to listen to the wide range of international perspectives offered by the fellows and to hear about the fantastic things they are doing to push life science forward.

I met and befriended people such as

Gaëlle Mainguy, a scientist in France who is the president of the World Academy of Young Scientists. This organization seeks to increase access to science for individuals in developing countries, an issue that has always been close to my heart. I was also able to discuss the possibility of future ventures in biotech with business consultants who advise life-science companies on financial matters.

Indeed, the tremendously positive and uplifting interactions I had with the fellows and our many discussions about future directions of research encouraged me to continue striving to improve global health through my ongoing research on cancer and AIDS. I left with a greater appreciation for my role as a member of an international community of young life scientists. I stay in frequent contact with several of the fellows, sharing experiences and advice, and hope to count them as future collaborators. I would recommend any recent PhD or young biotech entrepreneur to attend this meeting and take part in the programme. ■

Tshaka Cunningham is a postdoctoral fellow at the National Cancer Institute in Bethesda, Maryland.

For more information about applying to be a BioVision.Nxt fellow, contact Abigail Gemo at abigail.gemo@biovision.org.

GRADUATE JOURNAL

PhD survival guide

If you type the words 'survival guide' into any Internet search engine you will find that among many web pages dedicated to surviving the wilderness or various travel destinations, pages about making it through graduate school occur perhaps all too frequently.

Survival guides on a given topic exist because people have attempted to 'survive' the associated challenge — and have overcome significant obstacles during their quest. Based on the number of web pages dedicated to surviving graduate school, finishing a PhD is no walk in the park.

Much like climbing a mountain peak, completing a PhD will always be strenuous but, depending on the weather, the difficulty in reaching the summit will vary. Before you set out, you should check the weather forecast and make sure that your guide knows the mountain — or you may suddenly find yourself on a slippery slope. You should also make sure that you get along with your guide and that the two of you understand each other. If not, it is likely to be a very long climb to the summit — or you might not reach it at all.

With graduation in sight, it is probably no coincidence that my adviser, who is a mountaineer, has invited me to join him on an expedition to climb the world's highest active volcano the month before my dissertation defence. ■

Andreas Andersson is final-year PhD student in oceanography at the University of Hawaii.

My morning glory

What a way to start the day.

David Marusek

When I rise in the morning, I can hardly wait to run out to the living room and shout: "My Morning Glory! My Morning Glory!" Then My Morning Glory spins up and says, "Good morning, sir! You're out of bed early today — well ahead of schedule. We're off to a *brilliant* start on a brand new day!"

This is my first kudo of the day. I pump my arm in the air and shout: "Yes!" On the media shelf, My Kudo Kounter is blinking: *Keep up the good work!* Then My Personal Trainer says: "Today is Tuesday, and we all know what Tuesday is — Nimble Knees Day!" So I place my hands on my knees and — slowly at first — rotate them clockwise, then

shower. As I shave, My Mirror scrolls text messages — *Looking good there, champ* and *Did we lose a few pounds?* My Closet picks out a dark suit and says, "A striped tie would be perfect for our afternoon HR meeting."

HR meeting? Suddenly, my guts clench up like a fist. I check My Calendar, and sure enough, My Annual Evaluation is today. This afternoon! Somehow I had managed to forget all about it. On the dresser, My Frown Jar says, "Uh-oh, someone owes me a dollar."

"Shut up!" I shout at it. "Shut up! Shut up, all of you!" I drop the clothes on the floor and collapse on the bed. This will be my first performance review since the merger announcement last

"That's good news, indeed. Time check — we're running late."

I sigh and get off the bed, finish dressing, and gather up My Things. "Good-bye," I say as I climb into the airlock.

"Farewell, sir," My Morning Glory replies. "Have a spectacular day and remember, My Happy Hour will be right here waiting for you when you return." I lock the inner hatch, and as the air is being exchanged, I strap on My Filter Mask and wait in front of the door. But when the door slides open, I simply can't force myself to egress. Tears well up in my eyes.

"Is there something else, sir?" My Morning Glory says.

"They're going to fire me. I know it."

In a little while, when still I don't move, My Morning Glory says, "Don't worry, sir. You're a survivor. You're a top performer. You're practically a Force of Nature. You'll do just fine."

"But I don't feel like a Force of Nature. What if I don't measure up? I can't do this. I want to stay home."

My Morning Glory tuts and says, "Tell me, what's the Third Rule on the Road to Success?"

"Third Rule? Baby steps. One step at a time."

"Exactly! And what is your next step? The HR meeting this afternoon, or something earlier?"

I check My Calendar. "I don't have anything earlier."

"Oh, no? What happens at ten o'clock?"

"That's My Morning Coffee Break."

"You are correct, sir!"

My Morning Coffee Break, of course. At ten I take My Morning Coffee Break, which I love almost as much as My Morning Glory.

"Forget all about this afternoon, sir, and focus on making it 'til ten. That's all you have to do. Now get out of here and show them what you're made of."

The moment I step across the threshold, the door slams and bolts behind me. I lift My Filter Mask to wipe my eyes. Then I straighten up and march resolutely into My Future.

Thank you, My Morning Glory.

I'd be lost without you.

When not working on his second novel, *Mind over Oship*, Alaskan author David Marusek is training his cabin to fetch him his slippers. His debut novel, *Counting Heads*, was released by Tor Books in November 2005.

JACEY



faster and faster until My Personal Trainer says "Reverse direction!" and I wobble to a halt and start rotating the other way.

Meanwhile, My Channel is downloading the headlines. The economy is looking up, and consumer confidence is high. We and Our Coalition Forces are winning all the wars. Global disasters during the past 24 hours: 0.

"Breakfast is served!" sings My Kitchen. Oatmeal with raisins, coffee with creamer, and a big smile of cantaloupe — yum! My Morning Glory says: "Time check — we've got time to burn!" So, I enjoy a second coffee and take an extended

quarter, and I'm not ready for it. My numbers are down; my management confidence is low.

My Apartment grows silent as all its tiny motors spin down. The room is so quiet I can hear the groan of the city through the wall. Finally, My Morning Glory says, "Is everything all right?"

"No! Everything is not all right. Everything is a freaking disaster. I'm as good as dead."

"In that case, don't move. I'm calling an ambulance."

"Ambulance?" I say, sitting up. "I don't need an ambulance."